4-20-2015 12:00 AM

# Towards Intensified Protein Refolding and Purification on Size Exlusion Chromatography: Fundamental Studies and Mathematical Modelling

Pegah Saremirad
*The University of Western Ontario*

Supervisor
Dr. Ajay Ray
*The University of Western Ontario*

Graduate Program in Chemical and Biochemical Engineering
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

# TOWARDS INTENSIFIED PROTEIN REFOLDING AND PURIFICATION ON SIZE EXCLUSION CHROMATOGRAPHY: FUNDAMENTAL STUDIES AND MATHEMATICAL MODELLING

(Thesis format: Integrated Article)

by

PEGAH SAREMIRAD

Graduate Program in Chemical and Biochemical Engineering

A thesis submitted in partial fulfillment
of the requirements for the degree of PhD

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

# Abstract

For industrial recombinant-protein processes, protein refolding and purification are crucial steps towards the recovery of considerable numbers of active and safe therapeutic products. In this thesis, intensification strategies for protein refolding and purification processes are explored. Development of an intensified process aims at simultaneous optimization of process performance indicators, namely: refolding yield, product purity, volumetric productivity and solvent consumption which in turn decrease the cost and time constraints to market.

The first strategy investigated related to multivariable experimental work using size exclusion chromatography (SEC) as a refolding method and a denatured/reduced model protein (lysozyme). SEC was selected due to its potential for refolding of higher concentrations of protein compared to conventional refolding methods used currently in industry. The investigated variables were protein loading concentration, refolding buffer composition including pH, sodium chloride salt and L-arginine, aggregation prevention additive, concentrations. The interplay of these process variables was studied and it was shown when L-arginine is used, over the experimental space, the effects of pH and protein loading concentration on refolding yield are insignificant. This observation introduced the possibility of manipulating pH in a wider range without concerns for protein aggregation; for instance, to adjust the redox potential of the buffer without the need for costly redox couple chemicals to assist reformation of disulfide bridges in oxidative refolding of the protein. The results also provide more experimental evidence on the mechanism of aggregation prevention by L-arginine. Secondly an experimentally-verified model of oxidative protein refolding on SEC was developed, with the goal of high-throughput process screening and optimization using the aforementioned model. Model development involved exploration of methods to find characteristic information on short-lived refolding kinetic species and lysozyme oxidative refolding kinetic schemes and constants under the two studied refolding environments, namely with and without L-arginine additive. It was shown that L-arginine prevents aggregation without considerable impact on the kinetics of lysozyme oxidative refolding.

Finally, SEC in a multi-column continuous simulated moving bed configuration (SMB-SEC) was evaluated to fully exploit the potential of SEC for intensified protein refolding and purification. This configuration offers several advantages compared to single-column operation, including increased productivity per unit mass of solid phase, lower solvent consumption, and less diluted products, provided that operation parameters are screened and tuned for simultaneous optimization of process performance indicators. In this phase of the project, the effect of scale-up was predicted and considered for modifying and utilizing single-column model towards design/operation of a SMB-SEC. This thesis presents a framework for protein refolding and purification process development and optimization, including reduced cost of chemicals, improving the refolding yield, high-throughput measurements of parameters and finding a suitable reaction scheme of refolding and aggregation for mathematical model development applicable to both single-column and multi-column continuous operations, and defining appropriate process performance indicators for optimized operation of SMB-SEC.

## Keywords

Inclusion-body-based protein; Intensified oxidative protein refolding; Mathematical modeling; Multi-column continuous simulated moving bed.

# Co-Authorship Statement

This thesis contains materials which have been published as well as material submitted for publication.

Chapter 2: **P. Saremirad**, J.A. Wood, Y. Zhang, A.K. Ray. Multi-variable operational characteristic studies of on-column oxidative protein refolding at high loading concentrations. *J. Chromatogr. A* 2014, 1359, 70–75.

Chapter 3: **P. Saremirad**, J.A. Wood, Y. Zhang, A.K. Ray. Oxidative protein refolding on size exclusion chromatography at high loading concentrations: fundamental studies and mathematical modeling. *J. Chromatogr. A* 2014, 1370, 145-155.

Chapter 4: **P. Saremirad**, J.A. Wood, Y. Zhang, A.K. Ray. Oxidative protein refolding on size exclusion chromatography: from batch single-column to multi-column counter-current continuous processing, Manuscript in preparation for submission.

The preparation of the manuscript was carried out by the author of this thesis. All papers were co-authored and reviewed by Jeffery Alan Wood, Yan Zhang and Ajay Kumar Ray.

# Acknowledgments

Firstly, I would like to thank my supervisors Dr. Ajay K.Ray and Dr. Yan Zhang for providing this collaboration opportunity, and their support and guidance throughout this project. I am very grateful for the time and energy they have dedicated to this work.

I would also like to acknowledge Dr. Jeffery A. Wood for his mentorship and being constant source of technical support and scientific feedback and Valerie Orr for sharing her extensive knowledge and expertise particularly in the area of analytical methods used in this project. Special thanks to Dr. Lars Rehmann for his time and stimulating discussions around various topics of my work and permission to use analytical lab facilities, Dr. Lars Konermann for productive and informative discussions and being an extraordinary teacher, and Dr. Greg Vilk for introducing me to his many knowledgeable contacts in medical science and biochemistry departments whenever I needed technical or scientific help. The training on good laboratory practices and maintenance of the lab equipments by Ms. Ying Zhang is greatly appreciated. My current and previous lab mates and colleagues, who are so many to list, thank you!

Finally, I would like to express my gratitude to my family; my sister, Paniz, for being on my side throughout these years, her support and patience, my parents and my brother for their unconditional support and love.

# Table of Contents

# List of Tables

# List of Figures

# List of Abbreviations and Symbols

ANOVA: Analysis of variance

BSA: Bovine Serum Albumin

CV: Column volumes

DTT: Dithiothreitol

EDTA: Ethylene Diamine Tetra Acetic acid

IBs: Inclusion-bodies

MAC: Matrix assisted chromatography

SEC: Size exclusion chromatography

SMB-SEC: Simulated moving bed size exclusion chromatography

UV: Ultraviolet

$Bc_{b,exp}$: Experimental batch buffer consumption (mL/mg)

$Bc_{c,pred}$: Predicted continuous buffer consumption (mL/mg)

$C_{b,i}$: Solute *i* liquid-bulk-phase concentration (mg/mL)

$C_{D,i}$: Concentration of solute $i$ in the refolding buffer (mg/mL)

$C_{eqS,i}$: Solid equilibrium concentration (mg/mL)

$C_{f,i}$: Solute *i* concentration in feed (mg/mL)

$C_{Ra}^{CSS}$: Average cyclic steady state concentration at raffinate (mg/mL)

$C_s$: Critical concentration (mg/mL)

$C_{s,i}$: Solute *i* solid-phase concentration (mg/mL)

$D_L$: Axial dispersion coefficient ($cm^2$/min)

$k_a$: Aggregation reaction rate (mL/mg min)

$k_{app}$: Apparent refolding reaction rate (1/min)

$K_{eq,i}$: Solute *i* solid-liquid phase equilibrium constant (-)

$k_{ov,i}$: Overall mass transfer coefficient (1/min)

$k_{sec,i}$: Average distribution coefficient (size exclusion capacity)

$L$: Loading concentration (mg/mL)

$L_C$: Column length (cm)

$M$: Mass (mg)

$m_j$: Flow rate ratio in zone $j$ (-)

N: Native protein

$N_t$: Total number of columns (-)

$P$: Phase ratio (-)

$P_b$: Batch purity (%)

$P_{c,pred}$: Predicted continuous product purity (%)

$Pe$: Peclet number (-)

$Pr_{b,exp}$: Experimental batch productivity (mg/mL h)

$Pr_{c,pred}$: Predicted continuous productivity (mg/mL h)

$Q_b$: Batch flow rate (mL/min)

$Q_D$: Buffer flow rate (mL/min)

$Q_{Ex}$: Extract flow rate (mL/min)

$Q_F$: Feed flow rate (mL/min)

$Q_j$: Internal flow rate in zone $j$

$Q_{Ra}$: Raffinate flow rate (mL/min)

$R_b$: Batch recovery (%)

$R_{c,pred}$: Predicted continuous protein recovery (-)

$R_p$: Particle radius (cm)

$r_{b,i}$: Consumption or production rate of component $i$ at bulk-liquid phase (mg/mL min)

$r_h$: Apparent radius (nm)

$r_{s,i}$: Consumption or production rate of solute $i$ at solid phase (mg/mL min)

$t$: Time (min)

$t_e$: Elution time (min)

$t_{pulse}$: Pulse time (min)

$t_s$: Switching time (min)

$u$: interstitial velocity (cm/min)

U: unfolded protein

$V_0$: Elution volume of thyroglobulin from bovine thyroid (mL)

$V_c$: Column volume (mL)

$V_{e,i}$ : Solute $i$ elution volume

$V_{inj}$: Injection volume (mL)

$V_t$: Total porosity volume of the column (mL)

$x$: Axial distance (cm)

$Y$: Batch yield (%)

$\varepsilon$: Total porosity fraction (-)

$\varepsilon_b$: Bed void volume fraction (-)

$\varepsilon_p$: Particle porosity (-)

# Chapter 1

# 1 Introduction

## 1.1 Recombinant-protein Technology

The advent of recombinant DNA-technology in the 1970s opened the possibility of engineering and expressing valuable protein/peptide products using a host cell. Recombinant human insulin and growth hormone were first introduced to the market in the early 1980s improving the availability, quality and safety aspects of these products which before were only available by tissue extraction from plants and animals [1]. The market for recombinant bio-pharmaceuticals has been growing rapidly standing at 30 products with an estimated market value of 50 to 60 billion USD in 2004 [2] and reaching to more than 151 unique products approved by the Food and Drug Administration (FDA) and/or European Medicines Agencies for different clinical indications by 2012 [3].

Manufacturing of the biopharmaceuticals, including the above mentioned products, involves drug discovery, expression of the desired product in the host cell (upstream), cell line production (midstream), recovery and purification of active and pure product (downstream) and drug formulation. These stages are integrated and it is important that all phases of process development be designed in tandem for instance, screening the drug candidates considering both efficacy and downstream processing requirements. Otherwise, downstream operations can account for over 70% of the production cost as they are greatly impacted by the presence of impurities and contaminants [4].

For recombinant-protein production, selection of host cells includes prokaryotes, yeast, filamentous fungi, insect and mammalian cells. About 30% of the recombinant-protein-therapeutics are expressed in *Escherichia coli* (*E.coli*) due to its well-characterized genetics, rapid growth and high level of expression, and low cost media; the level of expressed product can reach up to 50% of the cell mass [1,3]. However, when *E.coli* is used to express a protein of interest, due to over-expression and lack of necessary condition (e.g. saturation of cell folding machinery), the protein molecules may not fold to their unique three-dimensional structure known as native functional conformation. The expressed unfolded/partially folded polypeptide chains are prone to intramolecular

interactions and aggregation. Particularly, when the expressed polypeptide chains contain cysteine residues, formation of intramolecular disulfide bridges under over-expressed condition promotes aggregation inside the cell [3,5]. These aggregates are called inclusion-bodies (IBs); they are insoluble and contain large amount of inactive protein. Table 1-1 shows some examples of therapeutics produced as IBs [1].

**Table 1-1.** Examples of IBs produced using recombinant DNA technology

| Molecule | Companies | Indication |
|---|---|---|
| rh Insulin and analogs | Eli Lilly, Aventis | Diabetes treatment |
| Interferon alfacon-1 | Valeant | Anti-viral<br>Anti-tumor<br>Anti-inflammatory |
| r Interferon β-1b | Schering AG, Chiron | |
| r Interferon γ-1b | Genentec<br>Intermune | |
| rh IL-1 receptor -antagonist | Amgen | Immunotherapeutic-agents |
| r IL-2 | Chiron | |
| r IL-2-diphtheria toxin fusion | Seragen / Ligand | |
| r IL-11 | Genetics Institute | |
| r Human growth hormone (r hGH) | Genentech, Eli Lilly, Pfizer, Schwartz-Pharma, Novo Nordisk | growth hormone deficiency |

## 1.2 Downstream Processing of Inclusion-body-based Protein

Figure 1-1 illustrates the common route for downstream processing of IBs. After cell harvest and disruption, IBs are easily separated from the impurities in the cell broth as they are usually heavier than other cell components resulting in low level of impurities (70-90% of total protein is often present in IBs) [3,5]. The separated IBs are commonly further processed by addition of chaotropic reagents (i.e. denaturing and reducing agents such as urea and DTT) or adjusting the pH of the buffer (pH induced solubilization) to break the intra-molecular bonds and dissolve aggregates. The latter method of solubilization is considered more economical but its applicability is limited and depends on the nature of the aggregates as not all IBs are dissolved by adjusting the pH [6]. Once aggregates are dissolved and unfolded protein monomer is obtained further downstream processing aims at recovering the protein monomer in its folded functional conformation (protein refolding) and eliminating remaining contaminates from the host cell if they are not separated during IB recovery as well as soluble aggregate impurities that might be formed during the process. Similar to *in-vivo* refolding of unfolded/partially folded polypeptide chains, the *in-vitro* refolding is also susceptible to aggregation and misfolding. In other word, the correct folding pathway competes with aggregation and misfolding. Separation of soluble aggregates is crucial as they can cause unwanted autoimmune responses [7–9]. These species are also linked to neurodegenerative diseases such as Alzheimer and Parkinson; ongoing research suggest that protein aggregation formation may be a symptom of the disease rather than cause of the disease [10]. Accordingly, if aggregates cannot be reduced to an acceptable level below the World Health Organization limits, a protein with therapeutic potential may be dropped from development.

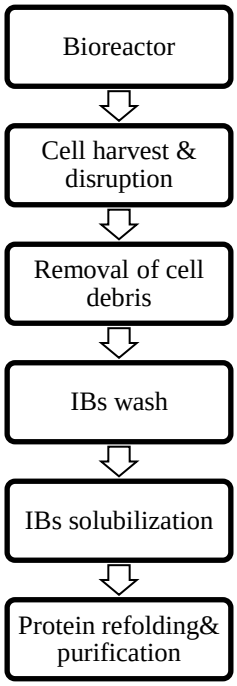


**Figure 1-1.** Downstream processing of inclusion bodies (IBs)

### 1.2.1 Protein Refolding Methods

Refolding strategies have involved two main categories to date, namely dilution and separation. In both cases reduced local concentration of denaturing and reducing agents surrounding unfolded protein molecule promotes formation of folded conformation. Batch dilution is considered the easiest method and used in industrial operation; however, the working concentrations are low to avoid aggregation formation which in turn results in higher buffer consumption, low volumetric productivity, handling large volumes and additional concentration steps [6]. Various matrix assisted chromatography (MAC) methods such as ion-exchange, hydrophobic, affinity and size exclusion chromatography (SEC), have been used at lab scale as separation methods [11–14]. MAC methods allow for higher protein concentrations to avoid aggregation due to spatial isolation of unfolded molecules while adsorbed on the matrix surface or gradual separation of chaotropic reagents from unfolded protein molecules. Although MAC techniques are also accompanied by dilution, considering the lower dilution factor and higher loading concentrations compared to dilution method, they result in more concentrated product. In

addition, the possibility of separation of impurities within the same unit holds a promise to eliminate the need for further processing to meet the standards in terms of impurity levels. Among different matrix assisted methods, SEC is a non-adsorptive technique and separation is achieved by different migration velocities according to the size of the species. The negligible interaction with the matrix may be advantageous compared to adsorptive methods. For instance, ion exchange chromatography may result in low recovery [15] unless multiple gradient schemes in terms of buffer pH, urea and salt concentrations are used [16]. Depending on properties of the adsorptive surface and protein, protein might remain unfolded and tend to aggregate while adsorbed on the surface. In these cases the highest productivity is achieved by low contact time with matrix and off-column refolding [17,18]. In contrast to hydrophobic chromatography, protein refolding on SEC requires refolding buffer containing low concentration of salt to prevent non-specific interactions [19], while high salt concentration requirement for protein binding in hydrophobic chromatography may be problematic when working with unfolded protein due to reduced solubility of these species compared to native conformation [5,20]. For these reasons, SEC has been considered for potential intensified protein refolding and purification and has been used at lab scale both in the form of batch single-column processing and multi-column continuous simulated moving bed operation (SMB-SEC) [6,19,21–27].

### 1.2.2 Fundamentals of Protein Folding/Refolding

In 1972 Christian B. Anfinsen (1972 Chemistry Nobel Prize) showed that native proteins can be unfolded using denaturing condition (e.g. breakage of disulfide bridges) but this unimolecular reaction can be reversible and native conformation of the protein is recovered under non-denaturing conditions [28]. Anfinsen's work resulted in the assumption that the native state is thermodynamically the most stable conformation under suitable conditions. However, this general acceptance was challenged by experimental protein aggregation studies. Baldwin et al. [29] observed that regardless of the sequence of the protein, when protein concentration surpassed a critical solubility concentration they aggregate to form amyloid fibril, a structure rich in β-sheet content, indicating that amyloid fibril is the most stable conformation . In the same work the cellular

concentration of the proteins under the study were compared to their equilibrium concentration and found to be higher than equilibrium concentrations. The question raised in this work is: how proteins have evolved to avoid aggregation during their normal life span *in-vivo?* One suggestion is that the native protein in cell is metastable compared to amyloid fibril and there is a large kinetic barrier between folded functional and aggregate states. Consequently, protein stability is controlled both thermodynamically and kinetically [29,30] yet there is no clear understanding that what causes kinetic stability of the protein inside the cell [31]. The overall stability of the native protein is the result of enthalpic (e.g. ionic bonds, hydrogen bonds and Van der Waals forces) and entropic contributions (release of disordered water molecules due to hydrophobic effect) that balance the large conformational entropy penalty associated with folding [32]. The native state is usually not much more stable than unfolded state and slight changes of the above mentioned interactions can induce unfolding such as temperature variation and chemical composition of the solvent.

Protein folding/refolding pathway has been classified to two main categories namely cooperative and non-cooperative. The former involves only two states (unfolded and folded) whereas the latter involves at least one intermediate state. Many proteins, especially small single domain species with less than 100 amino acids, fold in a two-state fashion [33]. However, larger proteins tend to deviate from two-state behavior. For example, the refolding of the model protein used in this work (lysozyme composed of 129 amino acids) involves a "collapse process" and at least one kinetic intermediate. There have been controversial approaches in describing collapse process (burst phase) [34]. The question is whether this phase represents a folding/refolding step or is a reflection of sudden change in the solvent properties and involves non-specific interactions. The latter description is supported by experimental evidence on observation of burst phase for proteins that cannot fold [35] whereas some research work debate specific nature of collapsed conformation (i.e. early intermediates) [36]. Further details regarding mechanism of lysozyme refolding and aggregation is provided in sections 3.2.2 and 4.2.3.

## 1.3 Research Contributions

Despite extensive experimental work and successful utilization of SEC at lab scale for protein refolding and purification, there is still work to be done to consider feasibility for industrial application of this technology. Major contributions of this work to the research field are outlined here and found in chapters 2, 3 and 4 of this thesis in more details.

**Chapter 2:** Multi-variable Operational Characteristic Studies of On-column Oxidative Protein Refolding at High Loading Concentrations [37]

The first contribution from this thesis relates to quantifying the interaction of L-arginine additive with refolding buffer pH and model protein loading concentration (lysozyme). A synthetic feed (denatured/reduced lysozyme obtained by addition of denaturing and reducing agents to naturally occurring lysozyme) was used throughout this work. The rationale for this approach was that regardless of the source of the protein (naturally occurring or recombinant), the refolding task involves dilution/separation of denaturing and reducing reagents from unfolded protein monomer. Lysozyme has been extensively studied as a model system for oxidative refolding due to importance of disulfide bond formation in considerable number of proteins.

It was shown that the effect of both pH and protein concentrations on the refolding yield in the presence of low concentrations of L-arginine (0.2 M) are trivial. In addition, in the presence of this additive and high protein loading concentrations (40 mg/mL) pore accessibility was reduced. Observations made in this study provided more experimental evidence towards the suggested mechanism of protein aggregation suppression by L-arginine and introduced the possibility of manipulating pH for adjusting the reducing potential of the buffer in oxidative refolding of the proteins, thereby potentially eliminating the need for an expensive redox couple commonly used in refolding practices.

**Chapter 3:** Oxidative Protein Refolding on Size Exclusion Chromatography at High Loading Concentrations: Fundamental Studies and Mathematical Modeling [38].

The second contribution is the development of an experimentally-verified mathematical model of oxidative protein refolding on SEC for fast screening of process parameters. A detailed kinetic scheme was investigated and various methods including quenched and equilibrium experiments as well as application of model compounds were introduced to obtain characteristic information on refolding kinetic species. It was demonstrated that characteristic information obtained from experiments on SEC may be erroneous due to non-specific interaction of species with the matrix. Kinetic studies showed an apparent two-state kinetic scheme including early intermediates and the native protein is significantly more accurate compared to commonly simplified mechanism including native-like intermediate to native protein. In addition under two explored chemical environments, namely with and without L-arginine, the mass transfer characteristic and kinetic studies showed that low concentration of L-arginine (0.2 M) increases the refolding yield of lysozyme with no considerable impact on the kinetics but did effect the mass transfer characteristics of lysozyme in SEC.

**Chapter 4:** Oxidative protein refolding on size exclusion chromatography: From batch single-column to multi-column counter-current continuous processing.

The third contribution from this thesis comes from introducing modifications into the model developed in the previous chapter in order to expand its applicability for prediction of protein behavior through columns connected in series in a SMB-SEC as a scale-up method. Aggregation was incorporated into the model to consider for higher local protein concentrations in SMB-SEC, due to the lower dilution factor in this system compared to the single-column operation. It was demonstrated that in a wide working concentration range the refolding/aggregation is best described as occurring when local protein concentration is equal or higher than a critical concentration (i.e. solubility of early intermediates), which can be measured. Examining the local denaturant concentration through the SMB-SEC confirmed the findings of previous research work that a model with constant kinetic and thermodynamic parameters may not provide an accurate estimate of early intermediate to native protein ratio at the product outlet. It was also shown that an unfolding reaction should be added to consider the higher local concentration of denaturant which occurs due to lower dilution factor compared to single

column. Based on predictions of the denaturant concentration at the product outlet it was concluded that the refolding reaction will continue off-column. Accordingly, process performance indicators were defined based on solubilized protein (early intermediates and functional native protein) which does not require differentiating between the solubilized conformations. This work demonstrated that a model with constant parameters is suitable for screening the operation parameters and optimization of the performance of a SMB-SEC unit if these effects are taken into account.

## 1.4 References


[1] K. Graumann, A. Premstaller, Manufacturing of recombinant therapeutic proteins in microbial systems, Biotechnol. J. 1 (2006) 164-186.

[2] F.R. Schmidt, Recombinant expression systems in the pharmaceutical industry, Appl. Microbiol. Biotechnol. 65 (2004) 363-372.

[3] C.J. Huang, H. Lin, X. Yang, Industrial production of recombinant therapeutics in Escherichia coli and its recent advancements, J. Ind. Microbiol. Biotechnol. 39 (2012) 383-399.

[4] R. Morenweiser, Downstream processing of viral vectors and vaccines, Gene Ther. 12 Suppl 1 (2005) S103-110.

[5] E.J. Freydell, Liquid Chromatography and Its Applicability to Protein Refolding, 2010.

[6] E.J. Freydell, L.A. van der Wielen, M.H. Eppink, M. Ottens, Techno-economic evaluation of an inclusion body solubilization and recombinant protein refolding process, Biotechnol. Prog. 27 (2011) 1315-1328.

[7] R. Jefferis, Aggregation, immune complexes and immunogenicity, MAbs 3 (2011) 503-504.

[8] J.G. Barnard, S. Singh, T.W.Randolph, J.F.Carpenter, Subvisible particle counting provides a sensitive method of detecting and quantifying aggregation of monoclonal antibody caused by freeze-thawing: Insights into the roles of particles in the protein aggregation pathway, J. Pharm. Sci. 100 (2010) 492-503.

[9] A.S. Rosenberg, Effects of protein aggregates: an immunologic perspective, AAPS J. 8 (2006) E501-507.

[10] C.A. Ross, M.A. Poirier, Protein aggregation and neurodegenerative disease, Nat. Med. 10 Suppl (2004) S10-17.

[11] L. Wang, X. Geng, Protein renaturation with simultaneous purification by protein folding liquid chromatography: recent developments, Amino Acids 46 (2014) 153-165.

[12] X. Geng, L. Wang, Liquid chromatography of recombinant proteins and protein drugs, J. Chromatogr. B Anal., Technol. Biomed. Life Sci. 866 (2008) 133-153.

[13] X. Geng, C. Wang, Protein folding liquid chromatography and its recent developments, J Chromatogr B Anal. Technol Biomed Life Sci 849 (2007) 69-80.

[14] A. Jungbauer, W. Kaar, Current status of technical protein refolding, J. Biotechnol. 128 (2007) 587-596.

[15] E.J. Freydell, L. van der Wielen, M. Eppink, M. Ottens, Ion-exchange chromatographic protein refolding, J. Chromatogr. A 1217 (2010) 7265-7274.

[16] M. Li, G. Zhang, Z. Su, Dual gradient ion-exchange chromatography improved refolding yield of lysozyme, J. Chromatogr. A 959 (2002) 113-120.

[17] E. Schmoeger, M. Wellhoefer, A. Dürauer, A. Jungbauer, R. Hahn, Matrix-assisted refolding of autoprotease fusion proteins on an ion exchange column: A kinetic investigation, J. Chromatogr. A 1217 (2010) 5950-5956.

[18] Y. Chen, S.S.J. Leong, Adsorptive refolding of a highly disulfide-bonded inclusion body protein using anion-exchange chromatography, J. Chromatogr. A 1216 (2009) 4877-4886.

[19] B.J. Park, C.H. Lee, S. Mun, Y.M. Koo, Novel application of simulated moving bed chromatography to protein refolding, Process Biochem. 41 (2006) 1072-1082.

[20] K.R. Reddy, H. Lilie, R. Rudolph, C. Lange, L-Arginine increases the solubility of unfolded species of hen egg white lysozyme, Protein Sci. 14 (2005) 929-935.

[21] M. Wellhoefer, W. Sprinzl, R. Hahn, A. Jungbauer, Continuous processing of recombinant proteins: Integration of refolding and purification using simulated moving bed size-exclusion chromatography with buffer recycling, J. Chromatogr. A 1337 (2014) 48-56.

[22] M. Wellhoefer, W. Sprinzl, R. Hahn, A. Jungbauer, Continuous processing of recombinant proteins: Integration of inclusion body solubilization and refolding using simulated moving bed size exclusion chromatography with buffer recycling, J. Chromatogr. A 1319 (2013) 107-117.

[23] Y. Chen, S.S.J. Leong, High productivity refolding of an inclusion body protein using pulsed-fed size exclusion chromatography, Process Biochem. 45 (2010) 1570-1576.

[24] E.J. Freydell, L.A. van der Wielen, M.H. Eppink, M. Ottens, Size-exclusion chromatographic protein refolding: fundamentals, modeling and operation, J. Chromatogr. A 1217 (2010) 7723-7737.

[25] E.J. Freydell, Y. Bulsink, S. van Hateren, L. van der Wielen, M. Eppink, M. Ottens, Size-exclusion simulated moving bed chromatographic protein refolding, Chem. Eng. Sci. 65 (2010) 4701-4713.

[26] H. Lanckriet, A.P. Middelberg, Continuous chromatographic protein refolding, J. Chromatogr. A 1022 (2004) 103-113.

[27] Z. Gu, Z. Su, J.C. Janson, Urea gradient size-exclusion chromatography enhanced the yield of lysozyme refolding, J. Chromatogr. A 918 (2001) 311-318.

[28] C.B. Anfinsen, Principles that govern the folding of protein chains, Science 181 (1973) 223-230.

[29] A.J. Baldwin, T.P.J. Knowles, G.G. Tartaglia, A.W. Fitzpatrick, G.L. Devlin, S.L. Shammas, C.A. Waudby, M.F. Mossuto, S. Meehan, S.L. Gras, J. Christodoulou, S.J. Anthony-Cahill, P.D. Barker, M. Vendruscolo, C.M. Dobson, Metastability of native proteins and the phenomenon of amyloid formation, J. Am. Chem. Soc. 133 (2011) 14160-14163.

[30] D. Thirumalai, G. Reddy, Protein thermodynamics: Are native proteins metastable?, Nat. Chem. 3 (2011) 910-911.

[31] K.A. Dill, J.L. MacCallum, The protein-folding problem, 50 years on, Science 338 (2012) 1042-1046.

[32] K.P. Murphy, Noncovalent forces in protein stability, Protein Stab. Fold., 1995, 1–30.

[33] H.S. Chung, W. a Eaton, a Single-molecule fluorescence probes dynamics of barrier crossing, Nature 502 (2013) 685-688.

[34] T.R. Sosnick, D. Barrick, The folding of single domain proteins-have we reached a consensus?, Curr. Opin. Struct. Biol. 21 (2011) 12-24.

[35] P.X. Qi, T.R. Sosnick, S.W. Englander, The burst phase in ribonuclease A folding and solvent dependence of the unfolded state, Nat. Struct. Biol. 5 (1998) 882-884.

[36] T. Kimura, S. Akiyama, T. Uzawa, K. Ishimori, I. Morishima, T. Fujisawa, S. Takahashi, Specifically collapsed intermediate in the early stage of the folding of ribonuclease A, J. Mol. Biol. 350 (2005) 349-362.

[37] P. Saremirad, J. A. Wood, Y. Zhang, A.K.Ray, Multi-variable operational characteristic studies of on-column oxidative protein refolding at high loading concentrations,J. Chromatogr. A 1359 (2014) 70-75.

[38] P. Saremirad, J.A. Wood, Y. Zhang, A.K. Ray, Oxidative protein refolding on size exclusion chromatography at high loading concentrations: Fundamental studies and mathematical modeling, J. Chromatogr. A 1370 (2014) 147-155.

# Chapter 2

**P. Saremirad**, J.A. Wood, Y. Zhang, A.K. Ray. Multi-variable operational characteristic studies of on-column oxidative protein refolding at high loading concentrations. *J. Chromatogr. A* 2014, 1359, 70–75

# 2 Multi-variable Operational Characteristic Studies of On-column Oxidative Protein Refolding at High Loading Concentrations

**Abstract**

Chromatographic-based protein refolding techniques have proven to be superior to conventional dilution refolding methods, due to the higher loading concentration and simultaneous purification. Among these techniques, Size Exclusion Chromatography (SEC) has in particular been demonstrated as an effective method for refolding of variety of proteins. To date existing studies of protein refolding at high concentrations (>1 mg/mL) in SEC have primarily been conducted as single factor studies, in which a single parameter is varied to assess impact on operating performance, which does not allow for determination of the interactions of different operating parameters and optimized operating conditions. In this work a multi-variable investigation of size exclusion protein refolding at high protein concentration using lysozyme as a model protein was performed, in order to quantify the interaction of factors and optimize performance. It was observed when L- arginine is used as an additive the refolding yield becomes independent of the protein concentration and refolding buffer pH, providing that a redox couple is used to assist the reformation of disulfide bridges. Furthermore, the pore accessibility for small molecules was reduced in the presence of this additive particularly at higher protein concentrations indicating slower removal of these molecules and a possible additional mechanism of aggregation prevention. Using the subsequent optimized refolding buffer, a refolding yield of more than 90% was obtained for up to 40 mg/mL loading concentration of lysozyme which has only been reported for a urea gradient SEC (8-2 M) with lower equilibration and elution flow rates due to high viscosity of buffer containing high concentrations of urea.

## 2.1 Introduction

Proteins are one of the most important biological compounds and beneficial to human health when used as therapeutic agents. Recombinant DNA technology continues to be one of the common methods in industry for production of many biopharmaceuticals,

including proteins [1]. In particular, *Escherichia coli* (*E. coli*) is one of the most used microbial expression systems in biotechnology due to its well characterized genetics, very high expression level and ease of manipulation [2]. One of the primary issues resulting from protein expression in *E. coli* is the formation of inactive protein aggregates (inclusion bodies). These aggregates require solubilisation through providing an environment for protein chains to unfold, which may be accomplished by using denaturing and reducing agents. After unfolding and aggregate collapse, the refolding of proteins into their compact structures is critical in order to restore biological activity and functionality. Refolding by dilution is commonly practiced in laboratories and industry due to its simplicity in design and operation [3–5]. However the correct protein folding pathway often competes with misfolding and aggregation, particularly at high concentrations which substantially reduces refolding yield. Furthermore, the presence of aggregates in the final product as impurities provides health concerns for utilization as therapeutics [6]. Consequently, the dilution technique has serious drawbacks during scale-up due to requiring low product concentrations and purity in addition to large process volumes, which necessitate additional cost-intensive post-refolding concentration and purification steps. These challenges limit high throughput production of therapeutic proteins and the speed with which new protein drugs can be brought to market [3].

Recently, chromatographic based refolding [3,7–9] has drawn great attention to address the challenges associated with product dilution, by facilitating spatial isolation of protein molecules and unfolding agents based on different affinity for solid phase or molecular size. These methods allow for protein refolding at higher concentrations and simultaneous protein purification due to reduced intramolecular interactions in adsorptive chromatography methods and gradual separation of protein and unfolding agents waves in non-adsorptive size exclusion chromatography (SEC). Among various chromatographic methods, SEC offers many advantages and has been widely used at lab scale for protein refolding in either batch or continuous mode [3,10–15]. The performance of SEC in terms of refolding yield, protein recovery and purity depends on many parameters, such as: protein structure, protein concentration, loading state of non-native protein (e.g. denatured, denatured and reduced), column packing specifications (e.g. material, particle size, pore size), refolding buffer composition including its pH,

redox potential, ionic strength and additives' concentrations [14–19]. However, the majority of research related to operational characteristics of SEC refolding pursued one-factor-at-a-time approach which cannot quantify the interactions of factors preventing determination of optimal operating conditions. In this work, a multi-variable study of key parameters on SEC refolding at high concentrations was carried out, using lysozyme as a model protein.

## 2.2 Design of Experiments

A suitable refolding buffer is particularly critical in refolding of proteins. The refolding buffer reported to give highest refolding yield for lysozyme is comprised of 0.1 M Tris, 1 mM EDTA, 2 M urea, 3 mM cysteine, 0.3 mM cystine or the same concentration of glutathione redox couple buffered at pH 8.1 [15]. In this work, the reported refolding buffer (0.1 M Tris, 1 mM EDTA, 2 M urea, 3 mM cysteine, 0.3 mM cysteine buffered at pH 8.1) was initially used to identify protein concentrations at which aggregates are formed. This is followed by a buffer optimization process to minimize the aggregation and increase the refolding yield. L- arginine is commonly used to increase the protein mass recovery in various liquid chromatography columns [20,21] and has been proven to be an effective aggregation suppressor due to its unique effects on protein association and folding [22–26]. Higher concentrations of L- arginine results in higher refolding yields, but it also slows down the rate of refolding [24,25]. Therefore, the concentration of L-arginine was selected as one of the key factors which affect the refolding yield of lysozyme by SEC. Apart from L- arginine concentration, refolding buffer pH, ionic strength and protein concentration are the other key factors which dictate the refolding yield of lysozyme [14,15,17,19]. A two-level full factorial design of experiment combined with replicated center point runs to test for curvature was executed in the current work to investigate the effect of the aforementioned operating parameters and the potential interactions between these factors. An empirical equation was developed to predict the refolding yield in the experimental space and search for optimum within the design space.

## 2.3 Materials and Methods

### 2.3.1 Chemicals

Reagent grade L- arginine and urea, Ethylene Diamine Tetra Acetic acid (EDTA), lysozyme from chicken egg white, trizma® base (Tris-base), L- cysteine, L- cystine, Bio-Ultra dithiothreitol (DTT) solution, *Micrococcus lysodeikticus*, potassium phosphate monobasic and BioXtra sodium chloride were purchased from Sigma-Aldrich, Canada. Red660™ protein assay reagent was purchased from G-Biosciences, USA. Superdex™75pg resin (24-44 µm) was purchased from GE healthcare, Canada.

### 2.3.2 Feed Preparation

Unfolding buffer (0.1 M Tris-base, 1 mM EDTA, 6 M urea and 32 mM DTT, pH 8.1) was used to prepare various concentrations of denatured and reduced lysozyme. The sample was incubated for 2-4 h at 37 °C to ensure loss of activity which was confirmed by enzymatic activity test as described below [15].

### 2.3.3 Refolding by Size Exclusion Column

XK16/40 column (GE healthcare, Canada) was packed with Superdex™75pg resin. The total volume of column was 44mL and the packing quality was tested by comparing the peak symmetry and number of theoretical plates per length of column with manufacturer recommended criteria using 2% (v/v) acetone injection. The packed column was installed on ÄKTA purifier 100, controlled by UNICORN 5.31 software equipped with online pH probe, UV detector and conductivity cell. The fractionation kit allows the collection of samples at desired volumes. The column was equilibrated with 2 column volumes (CV) of refolding buffer prior to protein injection. After equilibration, 0.5 mL of denatured and reduced lysozyme was injected and eluted for 1.5 CV with refolding buffer at 1 mL/min flow rate. During elution fractions of 7 mL were collected and stored at 4 °C before analysis which was conducted in less than 24 h. The stability of samples during storage was tested by comparing the enzymatic activity of samples analysed immediately and stored ones which showed no significant difference [27]. The fractions were pooled to

measure total protein recovery ($R_b\%$), refolding yield ($Y\%$) and purity ($P_b\%$) as defined below.

$$R_b = \frac{M_{total}}{LV_{inj}} \quad (2\text{-}1)$$

$$Y = \frac{M_{native}}{LV_{inj}} \quad (2\text{-}2)$$

$$P_b = \frac{M_{native}^{mono}}{M_{total}^{mono}} \quad (2\text{-}3)$$

where $M_{total}$ and $M_{native}$ are total protein and equivalent native protein mass collected in pooled fractions associated with either all forms of protein or protein monomer which were measured by total protein and enzymatic activity assays as described in analytical methods, $V_{inj}$ is injection volume, $L$ is lysozyme loading concentration.

The column was then washed with 2 CVs de-ionized water after elution. In case of in-column protein precipitation and flow blockage, the column was washed with 6 M urea, 32 mM DTT, 0.1 M Tris-base buffered at pH 8.1 at very low flow rates (< 0.2 mL/min) to dissolve the precipitated aggregates.

All the buffers were prepared fresh using ultra-pure water (Barnstead easy-pure RODI equipped with 0.2 µm filter, Fisher Scientific), filtered again with a 0.2 µm membrane and de-gassed prior to use.

### 2.3.4 Analytical Methods

#### 2.3.4.1 UV Absorbance

The lysozyme powder was dissolved in 0.1 M potassium phosphate and 0.15 M NaCl buffer (pH 7) and the lysozyme content was determined by ultraviolet (UV) absorption spectroscopy (Shimadzu UV-3600) at 280 nm using extinction coefficient of 2.63 mL/mg/cm. The feed concentration (denatured and reduced lysozyme) was confirmed

using extinction coefficient of 2.37 mL/mg/cm. Feed samples were diluted in 0.1 M acetic acid [27].

#### 2.3.4.2 Enzymatic Activity

In order to determine the concentration of equivalent native protein (activity recovery) in pooled fractions, the enzymatic activity of samples were compared with enzymatic activity of standard protein samples prepared in the same buffer as used for on-column refolding. The enzymatic activity was measured by recording the linear decrease of cell suspension absorbance (0.15 mg/mL *Micrococcus Lysodeikticus* in 0.1 M potassium phosphate buffer, pH 7) at 450 nm for 40 s [15], with all measurements in triplicate.

#### 2.3.4.3 Total Protein Concentration

The total protein concentrations in pooled fractions were determined using Red 660™ protein assay in which 50 µL protein samples was transferred to a test tube. 1 mL of reagent was added and mixed. The absorbance of the mixture was measured at 660 nm after 5 min. The total protein concentrations were calculated by comparing the sample and standard absorbance.

## 2.4 Results and Discussions

### 2.4.1 SEC Refolding of Lysozyme

Figure 2-1 (A and B) illustrate the chromatograms of SEC refolding of lysozyme at various protein concentrations (5-40 mg/mL) using refolding buffer of 0.1 M Tris, 1 mM EDTA, 2 M urea, 3 mM cysteine, 0.3 mM cystine. The fractions associated with aggregates and protein monomer were pooled to measure total protein recovery and the protein monomer fraction pool was used to measure purity and refolding yield as defined earlier, with the results summarized in Table 2-1. As expected, increasing protein concentration decreased the refolding yield due to the aggregation. However, under the current operating conditions (e.g. gel pore size, elution flow rate, refolding buffer composition), the aggregates are soluble and completely separated from protein monomer

in SEC column based on size, with larger aggregate molecules eluting before monomers. Therefore, close to 100% purity of protein monomer was attainable, highlighting the advantage of SEC for simultaneous protein refolding and purification after refolding. The last peak in all chromatograms corresponds to DTT, which due to low molecular weight can penetrate deeper in the gel pores.

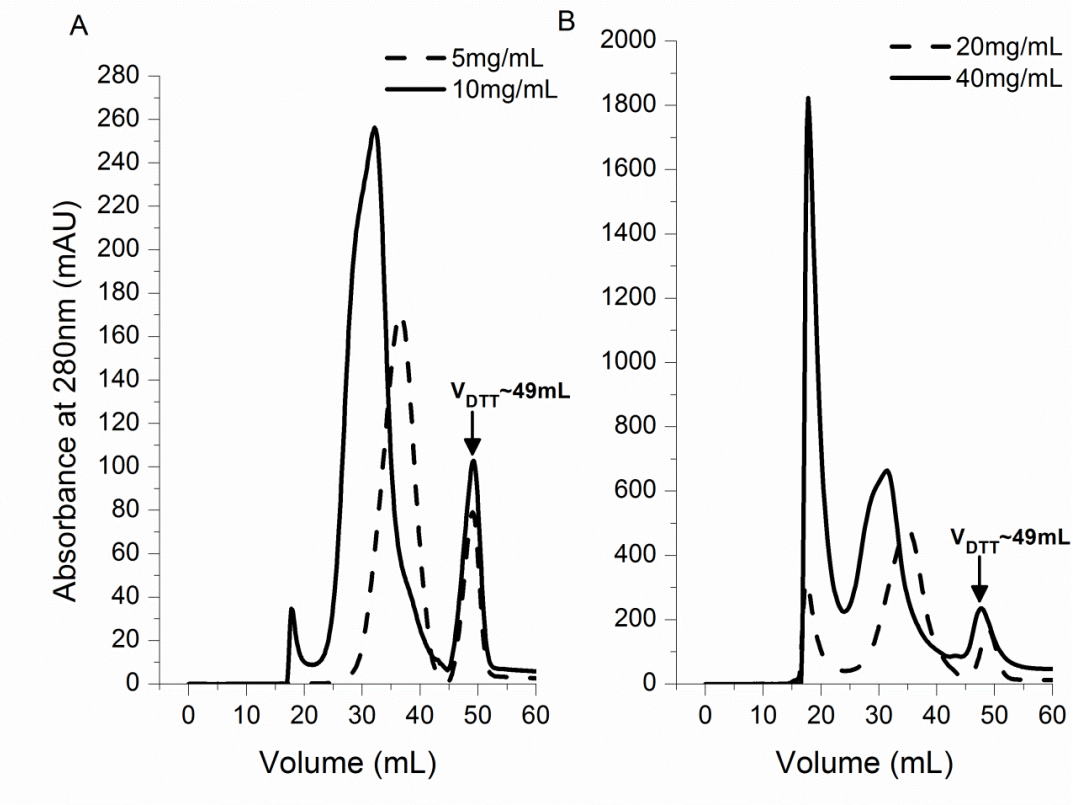


**Figure 2-1.** Chromatograms of SEC refolding of lysozyme (A) at low loading concentrations, (B) at high loading concentrations. Peaks from left to right: aggregate(s), protein monomer and DTT.

**Table 2-1.** SEC operating conditions, total protein recovery ($R_b$%), refolding yield ($Y$%) and purity ($P_b$%)

| $L$ (mg/ml) | $M_{inj}$ (mg) | $M_{total}$ (mg) | $M_{native}$ (mg) | $R_b$ (%) | $P_b$ (%) | $Y$ (%) |
|---|---|---|---|---|---|---|
| 4.5 | 2.25 | 2.3 | 2.48 | 100 | 100 | 100 |
| 9.05 | 4.52 | 4.47 | 4.3 | 98 | 100 | 95 |
| 18.26 | 9.13 | 9.56 | 7.7 | 100 | 100 | 84 |
| 36.03 | 18.01 | 16.1 | 11.37 | 89 | 95 | 63 |

### 2.4.2 Multi-variable Studies of Refolding of Lysozyme at High Concentration

A polynomial with linear and interaction terms was fit to measured experimental conditions to correlate the refolding yield (Y%) to lysozyme protein concentration (L) (20-40 mg/mL), refolding buffer pH (B) (8.1-9.5), NaCl salt (S) (0-0.2 M), L- arginine (A) (0-0.2 M) concentrations and their significant interactions. Table 2-2 reports the model coefficients, along with 95% confidence interval.

The analysis of variance (ANOVA) of the developed empirical fit showed the overall probability of 0.0001 which translates to 99.99% confidence that the coefficients are not zero. The model terms have p-values smaller than 0.05 indicating their significance within 95% confidence interval. Probability plots showed normal distribution of error and, as shown in Figure 2-2, good agreement between predicted and experimental results is achieved using this model for refolding yield >10%.

**Table 2-2.** Estimated empirical fit parameters and their confidence interval

| Parameter | Fit Coefficient | Relative Confidence- interval (%) |
|---|---|---|
| Intercept | 54.55 | 7.06 |
| *L* | -13.78 | 29.95 |
| *B* | -10.53 | 39.59 |
| *S* | -18.16 | 22.97 |
| *A* | 11.21 | 37.19 |
| *LS* | -5.84 | 70.69 |
| *LA* | 5.78 | 71.38 |
| *BA* | 8.28 | 50.35 |
| *SA* | -10.59 | 39.38 |

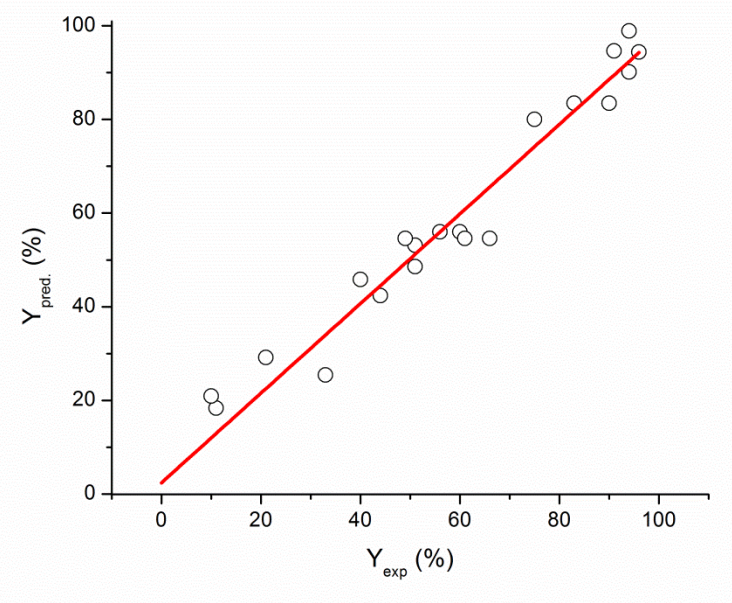


**Figure 2-2.** Experimental and predicted SEC refolding yield. The predicted SEC refolding yield was obtained using empirical fit.

The fit coefficients for individual parameters suggest that higher refolding yields are attained at low protein concentration, salt concentrations, refolding buffer pH and high L-arginine concentrations. However, as Figure 2-3 (A and B) shows when 0.2 M L- arginine is used as an additive in the refolding buffer protein concentration and refolding buffer pH demonstrated insignificant effect on the refolding yield in the experimental range tested in this work.

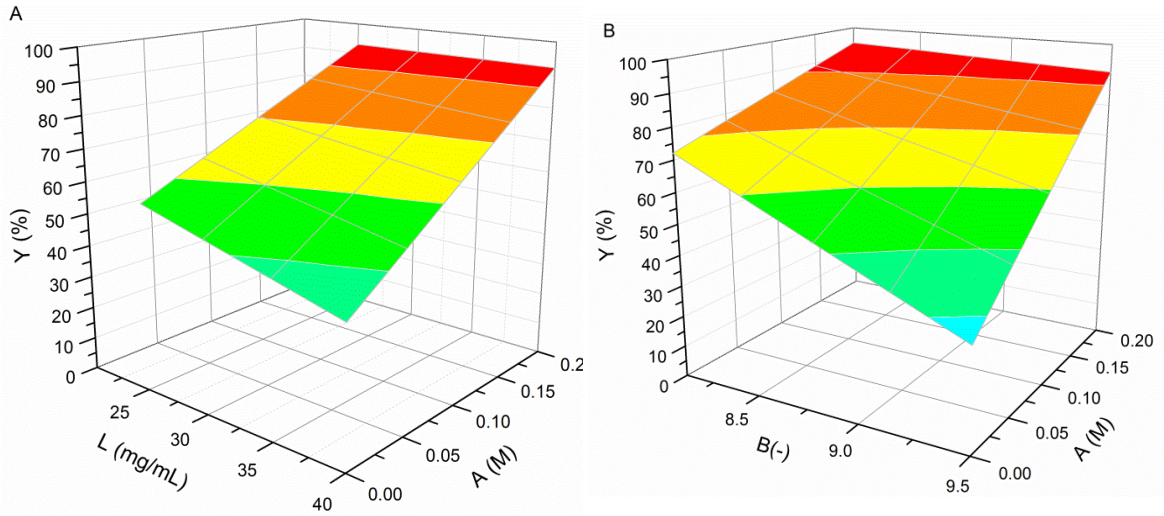


**Figure 2-3.** 3D surfaces generated using empirical fit; (A) pH = 8.8, no salt, redox couple (3 mM cysteine, 0.3 mM cystine) (B) Concentration = 30 mg/mL, no salt, redox couple (3 mM cysteine, 0.3 mM cystine). Y: refolding yield, L: lysozyme concentration, B: refolding buffer pH and A: L-arginine concentration.

High protein concentration can reduce the refolding yield due to increased intra-molecular interactions and protein aggregation. The refolding buffer pH influences the ionic interactions and redox potential of the system and is recognized as one of the most influential parameters in protein refolding. For lysozyme refolding, working at pH 9.5 decreased the refolding yield compared to pH 8.1. Although negligible amount of solubilized aggregates were formed at pH 9.5 (Figure 2-4), aggregation was observed outside the column in the fractions associated with protein monomers. Working at pH closer to protein isoelectric point (11.35) presumably reduces electrostatic repulsion and protein solubility, which can result in the aggregation of unfolded, intermediate or native species.

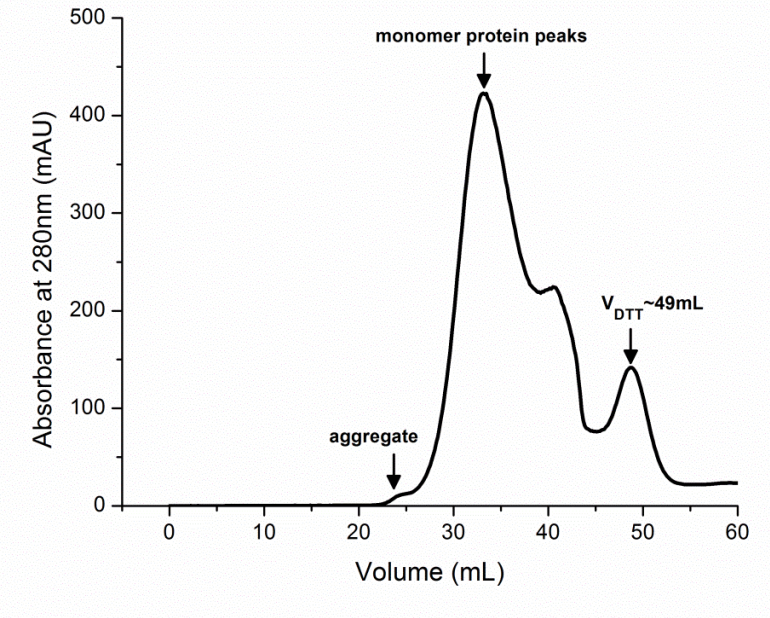


**Figure 2-4.** Chromatograms of SEC refolding of lysozyme at 20 mg/mL loading concentration and pH 9.5.

However, when L- arginine is used, formation of protein-associated arginine clusters "crowds out" the protein molecules and prevents self-association and aggregation [22,23]. As shown in Figure 2-5 (A and B) no aggregates were formed independent of lysozyme loading concentration and refolding buffer pH .It was hypothesized that the overlapping peak eluting before refolded lysozyme is an ensemble of early kinetic intermediates with only a fraction of correct disulfide bridges, little structure and buried tryptophan residues [28]. Refolding without a redox couple results in the recovery of species with similar characteristics by preventing rearrangement and formation of correct disulfide bonds [29]. Therefore, the refolding was also conducted without redox couple to gain more information on the species eluting before refolded protein. As Figure 2-5 (C) illustrates, at pH 8.1 no refolded lysozyme was recovered when the redox couple was removed and the eluting peak has the same retention volume as the first peak in the run with the redox couple (Figure 2-5B). The enzymatic activity test of the protein pool also confirmed no activity recovery. The higher redox potential at pH 9.5, allows for the refolding to proceed even without redox couple and self-association is prevented due to the presence of L- arginine. However, the reaction is slightly slower and the first eluting

peak is more populated. Nevertheless, its elution volume is identical to the experiment with the redox couple (Figure 2-5B). In both cases early kinetic intermediates were stable in the time scale of chromatography experiments and only partially precipitated outside the column overnight during storage. This is of particular interest when information about the size of kinetic intermediate is required for undertaking modelling and scale up.

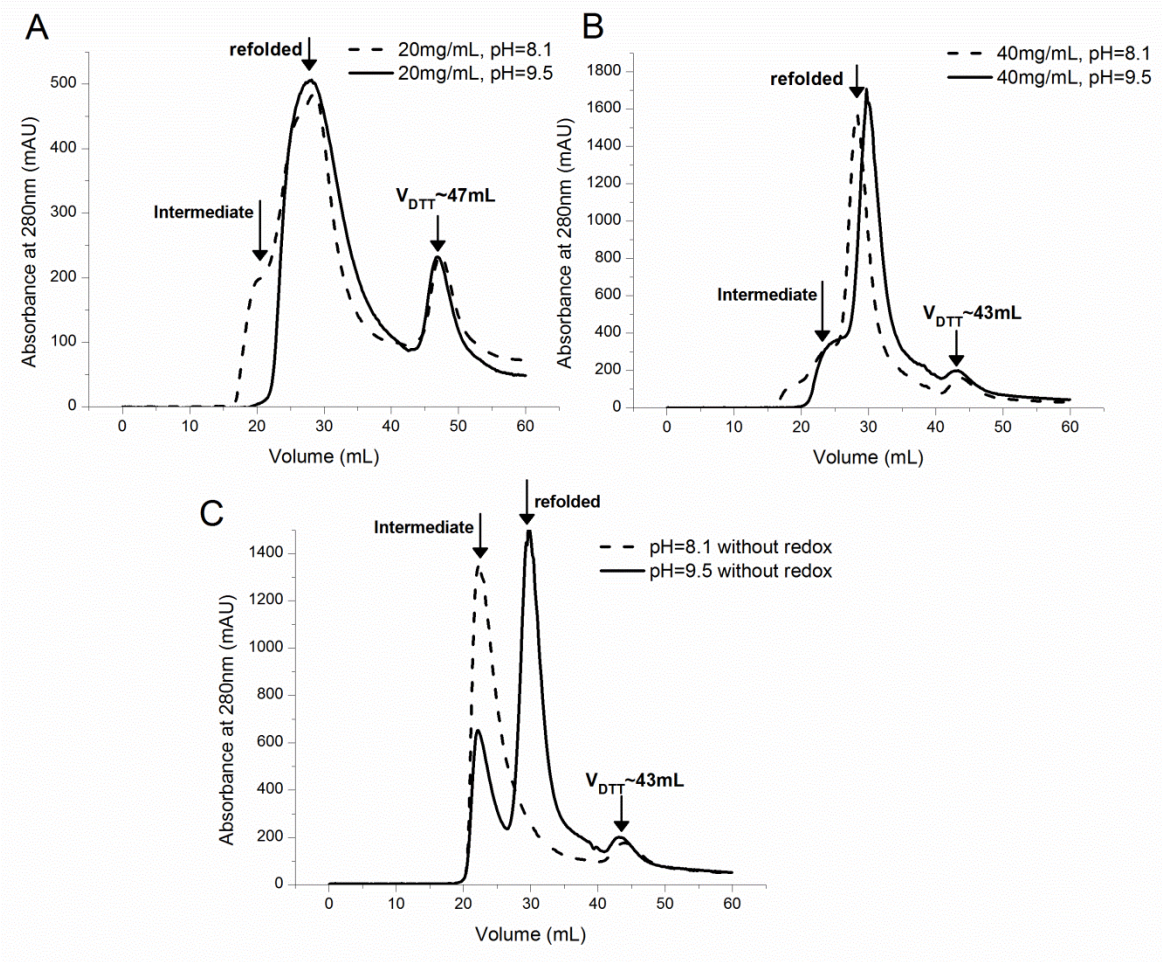


**Figure 2-5.** Effect of L-arginine using refolding buffer pH 8.1 and 9.5. (A) loading concentration of 20 mg/mL, (B) loading concentration of 40 mg/mL, and (C) loading concentration of 40 mg/mL without redox couple. The reduced pore accessibility due to presence of L-arginine is evident from DTT elution volume.

Another interesting observation of this work was faster elution of small molecules of urea and DTT when L- arginine is added to the refolding buffer particularly at high protein concentrations. DTT elutes at 48.7±0.5 mL when no arginine is used regardless of protein

concentration (Figure 2-1A, B and Figure 2-4). However, its elution volume is influenced by protein concentration in the presence of L- arginine eluting at 47±0.2 mL and 43.4±0.2 mL for loading concentrations of 20 mg/mL and 40 mg/mL respectively (Figure 2-5A, B and C). Forrer et al. [30] demonstrated that the accessible porosity of an ion-exchange material is reduced at high protein concentrations, due to the increased protein binding to the resin surfaces. For the current study in SEC column, reduced pore accessibility for small molecules may be responsible for their faster elution due to formation of protein-associated arginine clusters with larger sizes compared to protein monomer. The reduced transport properties for denaturing and reducing molecules achieved by the application of L- arginine at high protein concentrations can prevent sudden removal of these molecules and protein aggregation [15,17], another mechanism by which L- arginine can affect the on-column refolding yields at high concentrations.

Adding NaCl to refolding buffer led to in-column precipitation, and in some cases, flow blockage. Flow blockage happened immediately after injection due to decreased solubility of unfolded lysozyme which is no longer protected by urea molecules after their immediate diffusion into the gel pores [18]. This is in contrast to lower loading concentrations in which similar concentration of NaCl is used and recommended to reduce the non-specific interactions in SEC columns [14]. As an alternative, optimized concentration of L- arginine can be used to reduce the non-specific interactions with the gel and solid surfaces during purification and storage [21,31].

Using the subsequent optimized condition above 90% refolding yield was obtained for concentration as high as 40 mg/mL. The highest reported activity recovery for high concentration lysozyme (30.1 mg/mL) refolding in SEC is 80% [17] achieved by urea gradient SEC in which linear decrease of urea concentration along the column (8 M to 2 M) prevents the aggregation. However, the challenge of column equilibration with a buffer containing 8 M urea is high pressure drop due to high viscosity of the buffer.

## 2.5 Conclusions

The multi-variable investigation of lysozyme refolding in SEC at high concentrations carried out in this work revealed: (1) insignificant effect of protein concentration and refolding buffer pH on refolding yield at the presence of L- arginine provided suitable redox couple is used; (2) reduced pore accessibility for small molecules of urea and DTT in SEC at the presence of L- arginine and high protein concentrations; (3) early kinetic intermediates were identified during refolding with L- arginine and found to be stable in the time scale of experiments in this work which allowed for size measurements in SEC.

The possibility of working at high pH values for proteins with basic isoelectric points without self-association and aggregation problem due to the presence of L- arginine might eliminate the need for addition of expensive redox couple. The reduced pore accessibility for denaturing and reducing molecules is an important factor in SEC refolding of protein as it prevents sudden removal of these agents and could be an additional mechanism by which L- arginine additive prevents aggregation at high protein concentrations. These findings also provide more experimental evidence on proposed mechanism of arginine and protein interactions.

## 2.6 References


[1] K. Graumann, A. Premstaller, Manufacturing of recombinant therapeutic proteins in microbial systems, Biotechnol. J. 1 (2006) 164-186.

[2] C.J. Huang, H. Lin, X. Yang, Industrial production of recombinant therapeutics in Escherichia coli and its recent advancements, J. Ind. Microbiol. Biotechnol. 39 (2012) 383-399.

[3] E.J. Freydell, L.A. van der Wielen, M.H. Eppink, M. Ottens, Techno-economic evaluation of an inclusion body solubilization and recombinant protein refolding process, Biotechnol. Prog. 27 (2011) 1315-1328.

[4] A. Jungbauer, W. Kaar, Current status of technical protein refolding, J. Biotechnol. 128 (2007) 587-596.

[5] E.D. Clark, Protein refolding for industrial processes, Curr. Opin. Biotechnol. 12 (2001) 202-207.

[6] H.C. Mahler, W. Friess, U. Grauschopf, S. Kiese, Protein aggregation: pathways, induction factors and analysis, J. Pharm. Sci. 98 (2009) 2909-2934.

[7] L. Wang, X. Geng, Protein renaturation with simultaneous purification by protein folding liquid chromatography: recent developments, Amino Acids 46 (2014) 153-165.

[8] X. Geng, L. Wang, Liquid chromatography of recombinant proteins and protein drugs, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 866 (2008) 133-153.

[9] C. Machold, R. Schlegl, W. Buchinger, A. Jungbauer, A. Matrix assisted refolding of proteins by ion exchange chromatography, J. Biotechnol. 117 (2005) 83-97.

[10] M. Wellhoefer, W. Sprinzl, R. Hahn, A. Jungbauer, Continuous processing of recombinant proteins: Integration of refolding and purification using simulated moving bed size-exclusion chromatography with buffer recycling, J. Chromatogr. A 1337 (2014) 48-56.

[11] Y. Chen, S.S.J. Leong, High productivity refolding of an inclusion body protein using pulsed-fed size exclusion chromatography, Process Biochem. 45 (2010) 1570-1576.

[12] E.J. Freydell, L.A. van der Wielen, M.H. Eppink, M. Ottens, Size-exclusion chromatographic protein refolding: fundamentals, modeling and operation, J. Chromatogr. A 1217 (2010) 7723-7737.

[13] E.J. Freydell, Y. Bulsink, S. van Hateren, L. van der Wielen, M. Eppink, M. Ottens, Size-exclusion simulated moving bed chromatographic protein refolding, Chem. Eng. Sci. 65 (2010) 4701-4713.

[14] B.J. Park, C.H. Lee, S. Mun, Y.M. Koo, Novel application of simulated moving bed chromatography to protein refolding, Process Biochem. 41 (2006) 1072-1082.

[15] H. Lanckriet, A.P. Middelberg, Continuous chromatographic protein refolding, J. Chromatogr. A 1022 (2004) 103-113.

[16] B. Batas, C. Schiraldi, J.B. Chaudhuri, Inclusion body purification and protein refolding using microfiltration and size exclusion chromatography, J. Biotechnol. 68 (1999) 149-158.

[17] Z. Gu, Z. Su, J.C. Janson, Urea gradient size-exclusion chromatography enhanced the yield of lysozyme refolding, J. Chromatogr. A 918 (2001) 311-318.

[18] B. Batas, H.R. Jones, J.B. Chaudhuri, Studies of the hydrodynamic volume changes that occur during refolding of lysozyme using size-exclusion chromatography, J. Chromatogr. A 766 (1997) 109-119.

[19] B. Batas, J.B. Chaudhuri, Protein refolding at high concentration using size-exclusion chromatography, Biotechnol. Bioeng. 50 (1996) 16-23.

[20] T. Arakawa, K. Tsumoto, K. Nagase, D. Ejima, The effects of arginine on protein binding and elution in hydrophobic interaction and ion-exchange chromatography, Protein Expr. Purif. 54 (2007) 110-116.

[21] D. Ejima, R. Yumioka, T. Arakawa, K. Tsumoto, Arginine as an effective additive in gel permeation chromatography, J. Chromatogr. A 1094 (2005) 49-55.

[22] V. Vagenende, A.X. Han, M. Mueller, B.L. Trout, Protein-associated cation clusters in aqueous arginine solutions and their effects on protein stability and size, ACS Chem. Biol. 8 (2013) 416-422.

[23] D. Shukla, B.L. Trout, Interaction of arginine with proteins and the mechanism by which it inhibits aggregation, J. Phys. Chem. B 114 (2010) 13426-13438.

[24] J. Chen, Y. Liu, X. Li, Y. Wang, H. Ding, G. Ma, Z. Su, Cooperative effects of urea and L-arginine on protein refolding, Protein Expr. Purif. 66 (2009) 82-90.

[25] K.R. Reddy, H. Lilie, R. Rudolph, C. Lange, L-Arginine increases the solubility of unfolded species of hen egg white lysozyme, Protein Sci. 14 (2005) 929-935.

[26] K. Tsumoto, M. Umetsu, I. Kumagai, D. Ejima, J.S. Philo, T. Arakawa, Role of arginine in protein refolding, solubilization, and purification, Biotechnol. Prog. 20 (2004) 1301-1308.

[27] A.M. Buswell, A.P. Middelberg, Critical analysis of lysozyme refolding kinetics, Biotechnol. Prog. 18 (2002) 470-475.

[28] R. Silvers, F. Sziegat, H. Tachibana, S. Segawa, S. Whittaker, U.L. Gunther, F. Gabel, J.R. Huang, M. Blackledge, J. Wirmer-Bartoschek, H. Schwalbe, Modulation of structure and dynamics by disulfide bond formation in unfolded states, J. Am. Chem. Soc. 134 (2012) 6846-6854.

[29] B. Raman, T. Ramakrishna, C.M. Rao, Refolding of denatured and denatured/reduced lysozyme at high concentrations, J. Biol. Chem. 271 (1996) 17067-17072.

[30] N. Forrer, O. Kartachova, A. Butté, M. Morbidelli, Investigation of the porosity variation during chromatographic experiments, Ind. Eng. Chem. Res. 47 (2008) 9133-9140.

[31] Y. Shikiya, S. Tomita, T. Arakawa, K. Shiraki, Arginine inhibits adsorption of proteins on polystyrene surface, PLoS One 8 (2013) e70762.

# Chapter 3

**P. Saremirad**, J.A. Wood, Y. Zhang, A.K. Ray. Oxidative protein refolding on size exclusion chromatography at high loading concentrations: fundamental studies and mathematical modeling. *J. Chromatogr. A* 2014, 1370, 145-155.

# 3 Oxidative Protein Refolding on Size Exclusion Chromatography at High Loading Concentrations: Fundamental Studies and Mathematical Modeling

**Abstract**

Size exclusion chromatography has been demonstrated as an effective method for refolding a variety of proteins. However, to date process development mainly relies on laboratory experimentation of individual factors. A robust model is essential for high-throughput process screening and optimization of systems to provide higher productivity and refolding yield. In this work, a detailed kinetic scheme of oxidative refolding of a model protein (lysozyme) has been investigated to predict the refolding results in SEC. Non-reactive native, quenched and equilibrium studies were conducted to obtain the model parameters for the species formed during refolding of denatured/reduced lysozyme. The model was tested in various operating conditions, such as: protein loading concentration, injection volume, flow rate and composition of refolding buffer with and without the use of L-arginine additive. An apparent two-state mechanism was found adequate to describe refolding of lysozyme on SEC for the operating condition tested in this work. Furthermore, using low concentration of L-arginine combined with urea as common aggregation suppressor additives showed insignificant change in kinetics of refolding of lysozyme on SEC. However, addition of L-arginine changed mass transfer properties of some of the species formed in refolding reaction which was considered in the model to accurately predict the result of refolding on SEC.

## 3.1 Introduction

The conventional method of protein refolding, namely batch dilution refolding, requires working at low protein concentrations in order to prevent aggregation, which in turn results in low productivity impeding high-throughput protein refolding and downstream processing for production of many bacterially expressed recombinant proteins [1]. Size Exclusion Chromatography (SEC)-based protein refolding addresses this issue to some

extent by facilitating gradual spatial isolation of protein molecules and unfolding agents, which can prevent aggregation and allow for application of higher protein loading concentrations and simultaneous purification compared with batch dilution refolding. For these reasons SEC has been widely used at lab scale for protein refolding in either batch or continuous mode using single or multiple column configurations (i.e. simulated moving bed) [1–9]. Existing work has also developed mathematical models for separation in SEC using various model proteins in their native forms [10,11]. For protein refolding in SEC (reaction-separation SEC), the refolding reaction is incorporated into the mathematical model as well as interactions of aggregating species in case they are formed under the operating conditions being studied and the model parameters must be obtained for kinetic species formed during these reactions [5,12]. Development of an experimentally verified and robust model helps minimize the number of screening experiments for a broader range of operation condition and is essential for systematic process optimization.

The mechanism of oxidative refolding reaction is commonly simplified to include only one intermediate involved in the rate limiting step based on dilution oxidative refolding kinetic data [13,14]. However, the ratio of kinetic constants of refolding steps is influenced by chemical composition of the environment [13,14] and this simplification may not be applicable in all cases. This is important in SEC as the refolding reaction is accompanied with protein size variation and a correct scheme of such variation along the column is necessary to accurately predict the elution profile of the protein. Furthermore, in case of aggregate(s) formation, the simplified mechanism is not able to capture the interactions of all intermediates that are prone to aggregation [5,14,15]. In this work, (1) a mathematical model for refolding of denatured/reduced lysozyme in SEC at high loading concentrations was developed to investigate a detailed reaction mechanism previously proposed for this protein which was selected as a model system due to importance of disulfide bond formation in considerable number of proteins [13]; (2) non-reactive native, quenched and equilibrium experiments were executed to find the model parameters for the final product and short-lived kinetic intermediates formed during unimolecular refolding reaction; (3) the model was tested by varying operating conditions, namely: protein loading concentration, injection volume, flow rate and composition of refolding

buffer (with and without L-arginine additive); and (4) the effect of low concentration (0.2 M) of L-arginine on mass transfer and kinetic parameters in SEC was studied. L-arginine has been extensively studied and used to increase the refolding yield in batch dilution refolding by suppressing aggregation [16–19] and protein mass recovery in chromatography methods by decreasing non-specific interaction with the matrix [20,21]. In terms of kinetics of refolding, Reddy et al. [19] have reported insignificant change in apparent kinetic constant of oxidative refolding of lysozyme by batch dilution using guanidinium chloride and low concentration of L-arginine (up to 0.5 M) to suppress aggregation. However in contrast to this, Chen et al. [18] observed a considerable decrease when both urea (2 M) and L-arginine (0.5 M) were used in batch dilution refolding of recombinant human granulocyte colony-stimulating factor. Furthermore, Vagenende et al. [16] illustrated protein destabilization at the presence of low concentration of L-arginine (<0.5 M) by differential scanning calorimetric method which might in turn result in reduced folding kinetics. To our knowledge there is no conclusive information available on the effect of L-arginine on mass transfer properties of various refolding species and the kinetics of this reaction in SEC, making it a topic worthy of investigation.

## 3.2 Modeling

The protein refolding in size exclusion column was modeled using the transport-dispersive and solid-film linear driving force models formulated from differential mass balances for solutes in the bulk-fluid phase and the particle-solid phase respectively [5]. The governing equations are

$$\frac{\partial C_{b,i}}{\partial t} = D_L \frac{\partial^2 C_{b,i}}{\partial x^2} - u \frac{\partial C_{b,i}}{\partial x} - P k_{ov,i}\left(C_{eqS,i} - C_{s,i}\right) + r_{b,i} \tag{3-1}$$

$$\frac{\partial C_{s,i}}{\partial t} = k_{ov}\left(C_{eqS,i} - C_{s,i}\right) + r_{s,i} \tag{3-2}$$

where $C_{b,i}$ and $C_{s,i}$ are the concentration of solute $i$ (unfolded, intermediates and native conformations) in bulk-fluid phase and solid phase respectively. $t$ is time, $x$ axial

distance along the column, $D_L$ axial dispersion coefficient, $u$ interstitial velocity, $k_{ov,i}$ solute overall mass transfer coefficient, $P$ phase ratio, $C_{eqS,i}$ the solid phase concentration in equilibrium with the bulk concentration. $r_{b,i}$ and $r_{s,i}$ are the net concentration change due to refolding reaction in bulk and solid phases with details presented in section 3.2.2.

The solute solid phase concentration in equilibrium with the bulk concentration was treated as a linear equilibrium relationship with a fixed equilibrium constant [10]:

$$C_{eqS,i} = K_{eq,i} C_{b,i} \tag{3-3}$$

where $K_{eq,i}$ is the equilibrium constant.

The boundary and initial conditions used to solve equations (3-1) and (3-2) are:

$$C_{b,i}(t, 0^-) = \begin{cases} C_{f,i} & 0 < t < t_{pulse} \\ 0 & t > t_{pulse} \end{cases} \tag{3-4a}$$

$$\frac{\partial C_{b,i}}{\partial x}(t, L_C) = 0 \tag{3-4b}$$

$$C_{b,i}(0, x) = 0 \tag{3-4c}$$

$$C_{s,i}(0, x) = 0 \tag{3-4d}$$

where $C_{f,i}$ is solute concentration in feed, $t_{pulse}$ is the duration of sample injection, and $L_C$ is the column length. The assumption that the sample is introduced into the column as a rectangular pulse of length $t_{pulse}$ may not be valid in most practical applications. However, as the injection time is very small compared to retention time, such simplification still seems to be applicable [22].

To solve the above system, the first and second spatial derivatives were discretized using fourth-order finite difference equations except for boundary points for which second order forward and backward finite difference approximations were used. The resulting method of lines system of ODEs was solved numerically in MATLAB by ode15s solver.

### 3.2.1 Determination of Model Parameters

The axial dispersion coefficient was calculated based on the definition of the Peclet number, which was estimated using the correlation of Chung and Wen for small Reynold numbers as follows [5]

$$D_L = \frac{uL_c}{Pe} \tag{3-5a}$$

$$Pe = \frac{0.1L_c}{R_p \varepsilon_b} \tag{3-5b}$$

where $R_p$ is particle radius and $\varepsilon_b$ is bed void volume fraction which was measured using thyroglobulin from bovine thyroid as a test probe

$$\varepsilon_b = \frac{V_0}{V_c} \tag{3-6}$$

where $V_0$ and $V_c$ are the elution volume of thyroglobulin from bovine thyroid and column volume respectively, correcting by subtracting the system volume such as tubing and valves. The phase ratio (void to non-void volume) is

$$P = \frac{1 - \varepsilon_b}{\varepsilon_b} \tag{3-7}$$

The mass transfer and equilibrium constants ($k_{ov,i}$ and $K_{eq,i}$) were found by minimization of the deviation of measured concentration vs. calculated in a least squares sense using fmincon function in MATLAB with an additional constraint for recovery ($R_{b,i} \leq 1$) which was defined as

$$f(x) = \sum_{j=1}^{j=n} (C_{b,i}^{exp,j} - C_{b,i}^{fit,j}(x))^2 \tag{3-8}$$

$$R_{b,i} = \frac{M_i}{C_{f,i} V_{pulse}} \tag{3-9}$$

where $x$ is a vector of mass transfer and equilibrium constants, $C_{b,i}^{exp}$ the vector of experimental solute concentration at the column outlet, $n$ the number of elements in concentration vector, $M_i$ sum of mass of solute collected in fractions and $V_{pulse}$ is injection volume.The initial guesses for these parameters were estimated as follows:

$$K_{eq,i} = k_{sec,i} = \frac{V_{e,i-} V_0}{V_t - V_0} \tag{3-10}$$

where $k_{sec,i}$ is the average distribution coefficient (size exclusion capacity) of solute, $V_{e,i}$ is its elution volume and $V_t$ is total porosity volume of the column determined by pulse injection of acetone.

Freydell et al. [5] have provided a complete list of references which report correlations for calculation of protein free diffusivity, pore diffusivity and film mass transfer coefficient in order to calculate $k_{ov}$ .The hydrodynamic radius of the solute is necessary to estimate the above parameters. In the same paper, they proved the adequacy of extended-Ogston model to correlate average distribution coefficient to the hydrodynamic radius of proteins in two different gel materials namely Superdex75 and Sephaseryl 100. In this work, the same model was used to construct a calibration curve for the size exclusion column and find the molecular size of various species. This method is an efficient way to measure molecular size of kinetic species involved in a refolding reaction.

Finally, the kinetic constants in the rate terms were determined by minimizing $f(x)$for recovered native protein elution profiles with constrain of refolding yield ($Y_b \leq 1$) which was defined as

$$Y_b = \frac{M_{native}}{C_{f,U}\, V_{pulse}} \tag{3-11}$$

where $M_{native}$ is sum of native protein mass collected in fractions and $C_{f,U}$ is unfolded lysozyme concentration in feed.

### 3.2.2 Reaction Scheme

An adequate kinetic scheme is required to predict the on-column refolding results. Deviation of lysozyme refolding, in low denaturant environment (e.g up to 1 M urea), from two-state kinetics behaviour (cooperative refolding) is evident and its refolding involves at least one intermediate both for denatured and denatured/reduced forms of the protein [13,23–26]. However, lysozyme refolding kinetic data has been interpreted in many different ways and more than one "pathway model" can fit the experimental results [23]. Some research work suggests the existence of parallel refolding pathways which is justified by heterogeneity of unfolded state and includes a fast direct refolding from unfolded to native state and a slow refolding through formation of intermediates. However, only small fractions of molecules seem to follow the direct pathway [13,24,25].

Nevertheless, the refolding of reduced and denatured lysozyme has been successfully modeled using a single dominant reaction pathway in batch and fed batch refolding [14,15]. The reaction mechanism is shown in Figure 3-1; it is suggested that early intermediates with only a fraction of disulfide bridges are rapidly formed during "collapse process" followed by formation of native-like intermediates and a slow conversion of these intermediates to the native state [13].

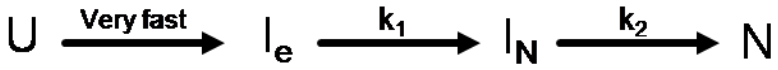


**Figure 3-1.** Lysozyme refolding kinetic scheme, U: unfolded protein, $I_e$: early intermediates, $I_N$: native-like intermediates and N: native protein.

According to this scheme the reaction rates for both bulk ($r_{b,i}$) and solid ($r_{s,i}$) phases in equations (3-1) and (3-2) are as following

$$r_{I_e} = -k_1 C_{I_e} \tag{3-12a}$$

$$r_{I_N} = k_1 C_{I_e} - k_2 C_{I_N} \tag{3-12b}$$

$$r_N = k_2 C_{I_N} \quad (3\text{-}12c)$$

Furthermore, formation of misfolded lysozyme with non-native contact, presumably incorrect disulfide bridges, has also been previously observed [8,27] and therefore might be included in the reaction mechanism. However, in this work very small fraction of misfolded protein was observed over the course of experiments. The misfolded species also slowly convert to native state [27] and therefore their formation was neglected for the purposes of this work.

## 3.3 Materials and Methods

### 3.3.1 Chemicals

Reagent grade Bovine Serum Albumin (BSA), L-arginine and urea, Ethylene Diamine Tetra Acetic acid (EDTA), lysozyme from chicken egg white, trizma® base (Tris-base), L-cysteine, L-cystine, BioUltradithiothreitol (DTT) solution, *Micrococcus lysodeikticus*, potassium phosphate monobasic, BioXtra sodium chloride, Trifluoric acetic acid (TFA) reagent plus grade and acetonitrile 0.1% TFA were purchased from Sigma-Aldrich, Canada. Red 660™ protein assay reagent was purchased from G-Biosciences, USA. Superdex™75pg resin (34 µm average particle size) was purchased from GE Healthcare, Canada.

### 3.3.2 Analytical Methods

#### 3.3.2.1 UV Absorption

Samples of native protein were prepared in 0.1 M phosphate solution buffered at pH 7 in order to determine the purity and lysozyme percentage of powder. The concentration of native lysozyme was determined with ultraviolet (UV) absorption spectroscopy (Shimadzu UV-3600) using extinction coefficient of 2.63 ml/mg/cm [27]. The concentration of native lysozyme standards prepared in refolding buffer was determined accordingly. The concentration of native protein in fractions collected during non-reactive native experiments was determined by using a Tecan M200 plate reader.

BSA standards were prepared in refolding buffer using 96% BSA powder and concentration of BSA in collected fractions were determined by comparing the fractions and standard protein samples UV absorbances using the plate reader.

#### 3.3.2.2 Native Lysozyme Concentration

A Vydac 214MS $C_4$ column (5 µm, 250 x 4.6 mm) was used on an Agilent HPLC system to separate native protein from other conformations and determine its concentration in collected samples during reactive experiments. A linear acetonitrile-water gradient with 0.1% (v/v) TFA starting at 25% acetonitrile increasing at 2.3%/min was used to elute the protein in 10min. The total solvent flow rate, column temperature and injection volume were set at 1 mL/min, 20 ºC and 50 µL respectively.

#### 3.3.2.3 Enzymatic Activity and Total Protein Concentration

The enzymatic activity of fractions was compared with activity of standard protein samples. The enzymatic activity of samples and standards were measured by recording the linear decrease of cell suspension absorbance (0.3 mg/mL *Micrococcus-lysodeikticus*in 0.1 M phosphate buffer, pH 7) at 450 nm after mixing using a microtiterplate reader after shaking for 40s [8].

The total protein concentrations in samples were determined using Red 660™ protein assay in which 10 µl protein samples was transferred to each well, 150 µL of reagent was added and mixed (6.5 mm circular shaker)using the plate reader, and absorbance of the mixture at 660 nm was measured after 5 min. The total protein concentrations were calculated by comparing the fraction and standard protein sample absorbance.

### 3.3.3 Feed Preparation

Unfolding buffer (0.1 M Tris-base, 1 mM EDTA, 6 M urea and 32 mM DTT, pH 8.1) was used to prepare various concentrations of denatured/reduced lysozyme (2-20 mg/mL). The excess of DTT was used to assure complete reduction of disulfide bridges. The sample was incubated for 2 h at 37 °C [8] and loss of native structure was confirmed by RP-HPLC analysis afterwards.

### 3.3.4 Experimental Set up

A XK16/40 column (GE healthcare, Canada) was packed with Superdex™75pg resin (column volume ~54 mL). The packed column was installed on ÄKTA purifier 100 controlled by UNICORN 5.31 software and equipped with online pH probe, UV detector and conductivity cell. The fractionation kit allows the collection of samples at desired volumes. Acetone pulse injection (2% (v/v)) was used to test the packing quality by comparing the peak symmetry and number of theoretical plates per length of column with manufacturer recommended criteria. Column void volume is commonly determined by blue dextran; however it was observed that blue dextran binds to superdex75pg and does not elute from the column despite using recommended concentration of salt in the mobile phase. Therefore, thyroglobulin from bovine thyroid (669 kDa) which is completely excluded from the macropores of Superdex75pg (fraction range of 3-70 kDa) was injected on the column to measure the void fraction.

### 3.3.5 SEC Non-reactive Experiments

Various concentrations and volumes of native lysozyme was injected on the column equilibrated with 2 column volumes (CV) of a refolding buffer identical to buffer used for reactive experiments. The protein was eluted using the same buffer and 1mL/min flow rate while 0.5 mL fractions were collected for UV absorbance analysis to determine the native protein concentration in each fraction. The model parameters for native species were fitted against these experimental results. Refolding buffer A was composed of 0.1 M Tris, 1 mM EDTA, 2 M urea, 0.2M L-arginine, 3 mM cysteine, 0.3 mM cysteine buffered at pH 8.1 while in refolding buffer B, L-arginine is replaced by 0.1 M NaCl to investigate the implication of removing L-arginine in terms of mass transfer parameters and kinetics of refolding. The addition of NaCl was necessary to prevent non-specific interactions with the column [7,20].

The redox couple was then removed from buffer A to quench the refolding at early intermediates [28,29] and fit the model parameters for these species. Fractions of 0.5 mL were collected and analyzed using total protein assay.

Native-like intermediates of lysozyme have been recovered in batch experiments by acid-quenching and HPLC separation of samples recovered at later stages of refolding [13]. Alternatively, it may be possible to gain information on the structure of kinetically short-lived intermediates by performing equilibrium experiments [23]. In the case of lysozyme, the characteristics of kinetic native-like intermediates formed in refolding of denatured and denature-reduced hen egg lysozyme are very similar [13] and equilibrium intermediate has been detected in unfolding experiments using urea where native-like intermediates showed the highest population at 4M urea [30]. Therefore equilibrium studies were performed to find the model parameters for native-like lysozyme. The column was equilibrated with 2 CVs of buffer C (0.1 M Tris, 1 mM EDTA, 4 M urea, 0.2 M L-arginine) and the same buffer was used to dissolve lysozyme which was incubated over night at 37 ºC. The protein was eluted using 0.5mL/min flow rate to avoid high pressure drop due to the high concentration of urea present in the equilibrium buffer C. Fractions of 0.5 mL were collected and analysed using total protein assay.

### 3.3.6 SEC Refolding- Reactive Experiments

The column was equilibrated with 2 CVs of the refolding buffer A and various volumes (0.5 and 1 mL) and concentrations (5, 10 and 20 mg/mL) of denatured/reduced lysozyme was injected and eluted using the same buffer and flow rate of 1mL/min with exception of one run where 0.5mL/min flow rate was used. The above loading concentration range was selected based on previous studies which demonstrated no aggregation formation within this range [31]. Fractions of 1mL were collected and analyzed using RP-HPLC to determine the concentration of correctly refolded protein. The enzymatic activity and total protein analysis were also carried out for some reactive experiments to gain more information on characteristics of recovered species.

For reactive experiments without L-arginine, the column was equilibrated with 2 CVs of the refolding buffer B and various volumes (0.5, 1 and 2 mL) of 2 and 5 mg/mL denatured/reduced lysozyme were injected and eluted with the same buffer. Low loading concentrations were selected to assure no aggregates were formed. Fractions of 1mL were again collected and analysed using RP-HPLC.

All the buffers were prepared fresh using ultra-pure water (Barnstead easy-pure RODI equipped with 0.2 µm filter, Fisher Scientific), filtered again with a 0.2 µm membrane and de-gassed prior to use. The column was washed with 2 CVs de-ionized water after final elution.

## 3.4 Results and Discussion

### 3.4.1 Non-reactive Experiments

The average experimental distribution coefficient ($k_{sec,N}$) and fitted values for the equilibrium solid phase concentration ($K_{eq,N}$) and the overall mass transfer coefficient ($k_{ov,N}$) of the native protein using buffers A and B are reported in Table 3-1. In both cases, $K_{eq,N}$ is less than $k_{sec,N}$. Figure 3-2 shows the satisfactory agreement between experimental and predicted results for native protein elution profiles. Reduced average distribution coefficient, equilibrium solid phase concentration and overall mass transfer coefficient when buffer A is used compared to buffer B, indicated a larger apparent radius of native protein in the presence of L-arginine. Vagenende et al. reported similar results and explained this observation based on formation of "dynamic arginine-protein associated clusters" extended from the surface of the protein [16]. As discussed earlier, the apparent radius was extrapolated using column calibration curve (Figure 3-3) based on extended-Ogston model. The lysozyme hydrodynamic radius measured in this work is larger than some reported values [30] as lysozyme swells in the presence of urea but is consistent with results reported for urea [32].

**Table 3-1.** Average distribution coefficient, fitted model parameters and apparent radius of kinetic species formed during lysozyme refolding. Solute ***i*** is native lysozyme for first and second row early and native-like intermediates for third and fourth row respectively.

| Buffer | $k_{sec,i}$ (-) | $K_{eq,i}$ (-) | $k_{ov,i}$ ($min^{-1}$) | Apparent $r_h$ (nm) |
|---|---|---|---|---|
| A | 0.55±0.01 | 0.44±0.003 | 17.29±1.99 | 1.8 |
| B | 0.58±0.01 | 0.49±0.002 | 21.61±1.55 | 1.7 |
| A and B | 0.08±0.01 | 0.08±0.004 | 2.69±0.45 | 3.5 |
| C | 0.47± 0.004 | 0.39±0.001 | 16.19±0.22 | 2 |

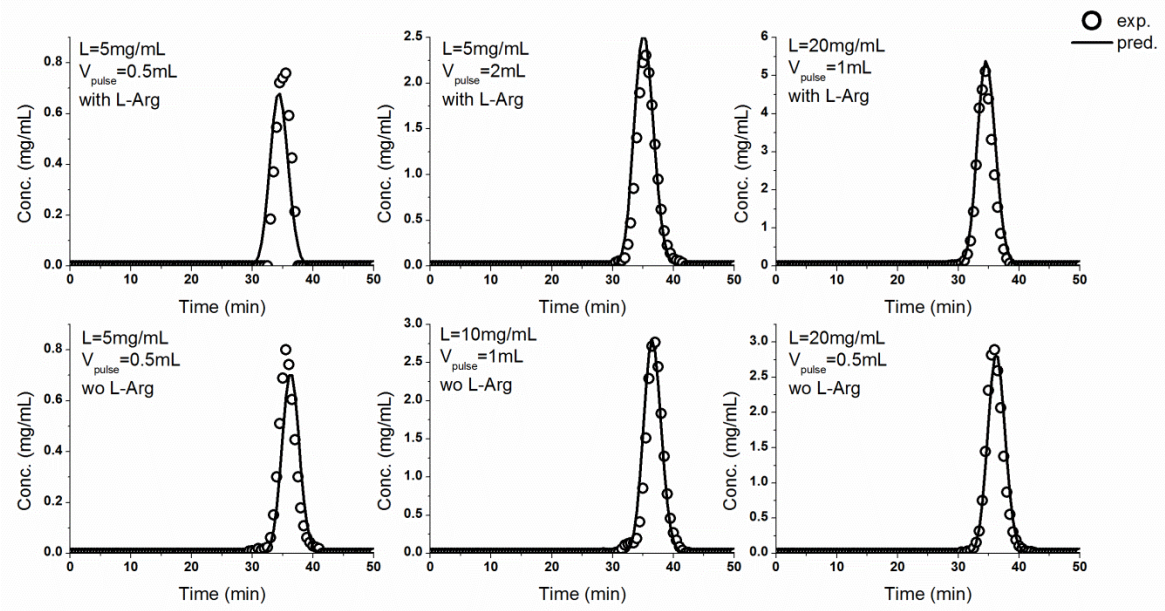


**Figure 3-2.** Experimental vs. predicted native protein elution profiles using buffer A (top row) and B (bottom row).

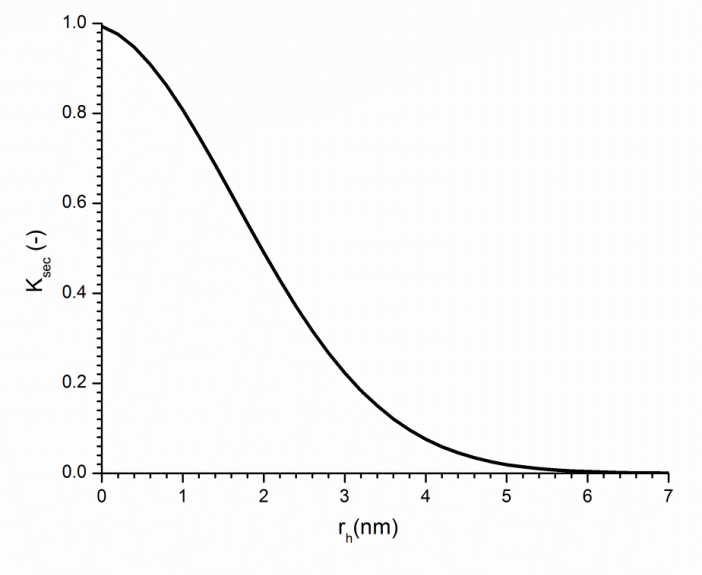


**Figure 3-3.** Size exclusion column calibration curve based on extended-Ogston model.

Quenched experiments were carried out to determine model parameters. Figure 3-4A shows the agreement between experimental and predicted elution profiles of early intermediates using fitted (least squares) values of 0.24±0.002 and 14.37±1.97 $min^{-1}$ respectively. However, early intermediates under quenching conditions showed a higher elution volume compared with reactive experiments (Figure 3-6A). This may be due to increased non-specific interaction at higher concentration of early intermediates with the gel in quenched experiments compared to short-lived early intermediates in reactive experiments. However, this is in contrast with that presented in our previous work [31] where the same elution volumes were observed and suggest that Superdex75pg surface properties changes over time well before expiry of the gel as indicated by manufacturer (~ 2 years ). Binding of blue dextran to the gel during void volume measurements (as mentioned before) was another indication of such change over time. The recovered protein still did not show any enzymatic activity and the florescence intensity was higher than native protein (data not shown). Since elution position of early intermediates was seemingly affected by interactions with the gel matrix during quenched experiments, the model parameters were fitted for BSA protein (66.5 kDa) elution profiles and used as a model compound (Figure 3-4B). BSA was used because it showed the same elution volume as early intermediates in reactive experiments suggesting the same extent of

exclusion within fraction range of the gel. As shown in Figure 4B, the elution volume of BSA was not influenced by the type of refolding buffer and the same model parameters were fitted for when buffer A or B were used as reported in Table 3-1.

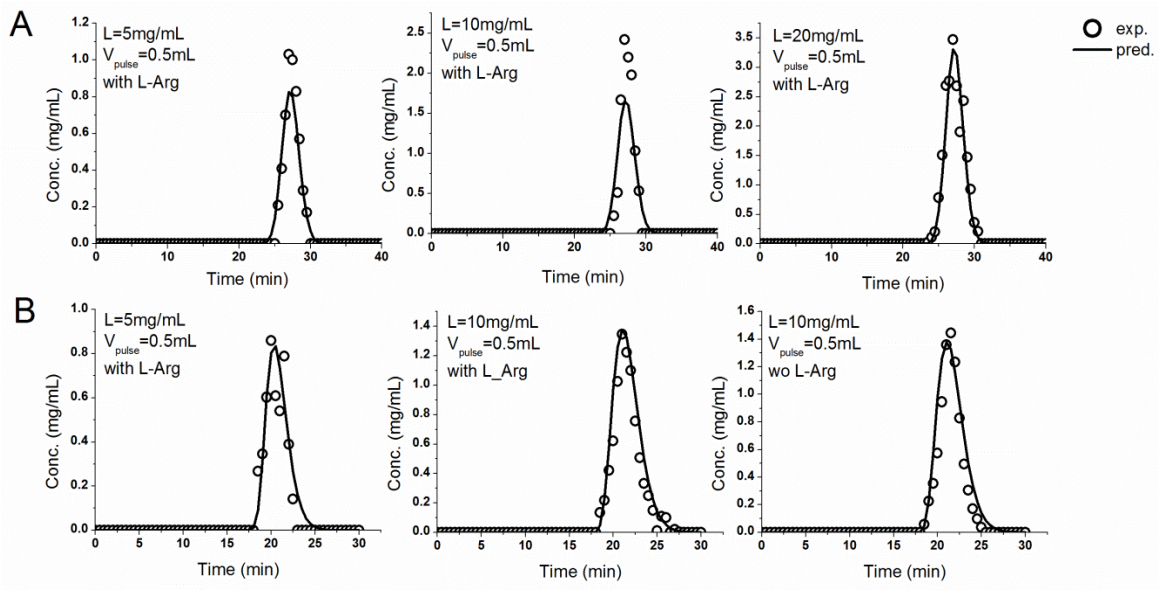


**Figure 3-4.** (A) Experimental vs. predicted early intermediate elution profiles using buffer A without redox couple to quench the reaction at early intermediates (B) Experimental vs. predicted BSA elution profile used as a model compound to find model parameters for early intermediates.

Figure 3-5 shows the elution profiles of lysozyme equilibrated in refolding buffer C. As described in section 3.3.5 the model parameters and size information for native-like intermediate were calculated using equilibrium studies (summarized in Table 3-1).

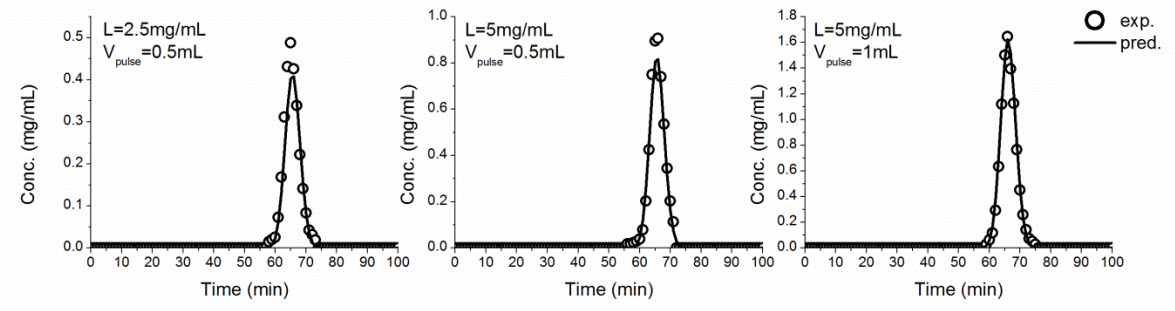


**Figure 3-5.** Experimental vs. predicted elution profiles of lysozyme equilibrated in buffer-C.

### 3.4.2 Reactive Experiments

Three protein conformations were recovered in reactive experiments using refolding buffer A, as shown in Figure 3-6A.

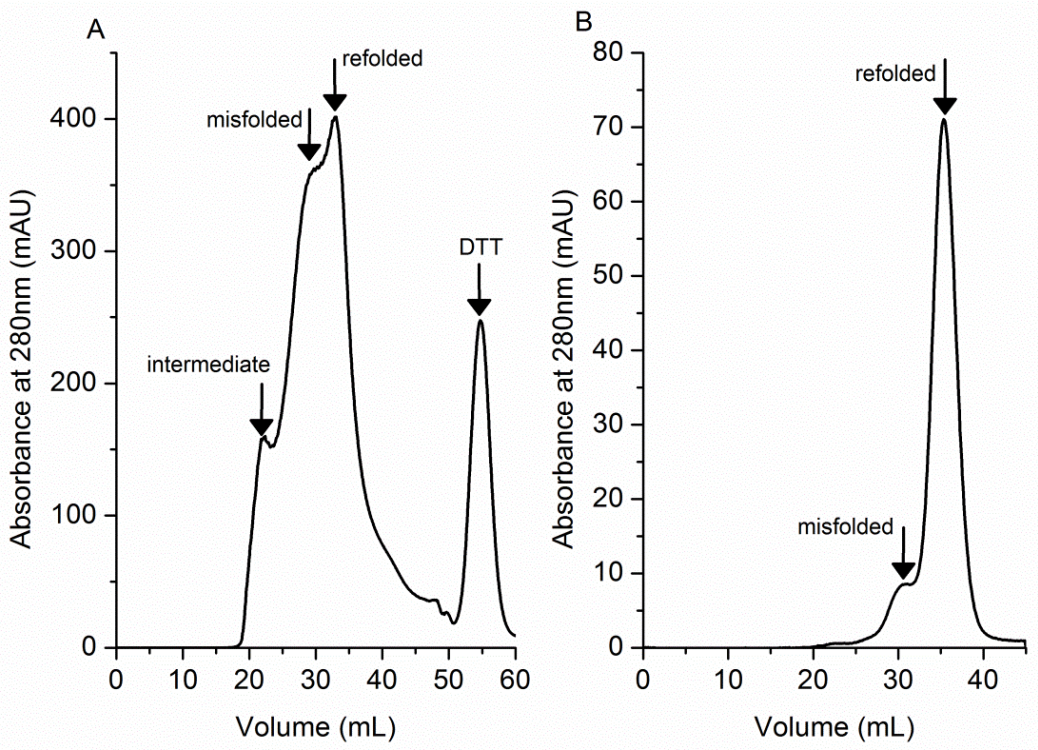


**Figure 3-6.** (A) Chromatogram of refolding of lysozyme using refolding buffer A (L=10 mg/mL, $V_{pulse}$=1 mL) and (B) Chromatogram of protein pool obtained from reactive experiment of (A).

The first peak is associated with early intermediates with no enzymatic activity. The second and third peaks showed similar enzymatic activity based on measurement with enzymatic activity and total protein assays. However, considering the elution profile of the recovered protein pool (Figure 3-6B), it is obvious that only one peak corresponds to native lysozyme eluting at the same volume as non-reactive native protein injections and the peak eluting before that was characterized as misfolded protein. It can be observed on Figure 3-6A that fractions containing early intermediates also contain misfolded protein (overlapping peaks) and quantification of either early intermediate or misfolded was not possible by HPLC analysis due to unavailability of standards for these species. The difference between total protein concentration and native protein is the summation of these two species and they cannot be quantified individually. This highlights the need for conducting quenched experiments or use of a model compound such as BSA to find model parameters to provide estimates for early intermediates.

A simplified reaction scheme considering native-like intermediate to native protein as the rate-limiting step was initially used to find the reaction rate constants and simulate the results of on-column refolding in this work. However, as shown in Figure 3-7 it was observed that this mechanism does not produce a satisfactory agreement between experimental and predicted values due to exclusion of early collapsed to native-like intermediate step from refolding reaction.

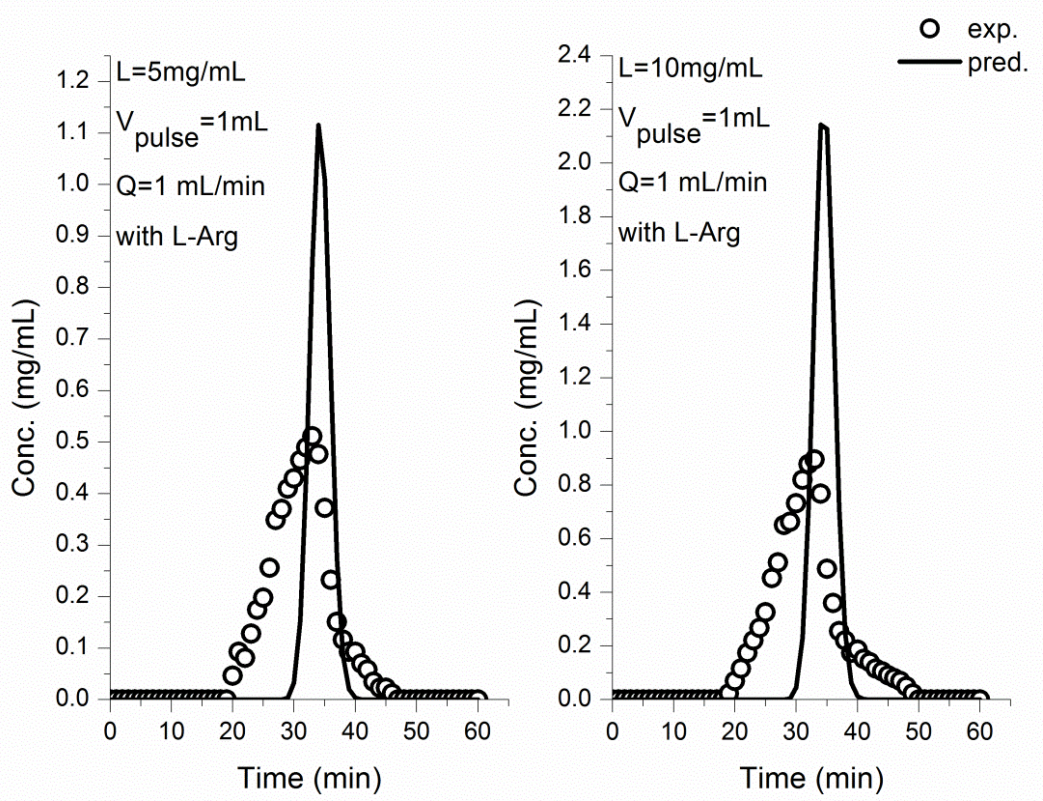


**Figure 3-7.** Experimental vs. predicted native protein elution profiles recovered during reactive experiments using a simplified reaction scheme from native like intermediate to native lysozyme.

On the other hand, an apparent two-state refolding representation of the reaction mechanism including early collapsed intermediates to native protein showed a satisfactory fit as shown in Figure 3-8. For one case of low flow rate (0.5 mL/min), the same mass transfer parameters as higher flow rate (1 mL/min) were used because estimated film and pore mass transfer resistances (determined using correlations) revealed the pore resistance as the major resistance component and this value is independent of velocity [10].

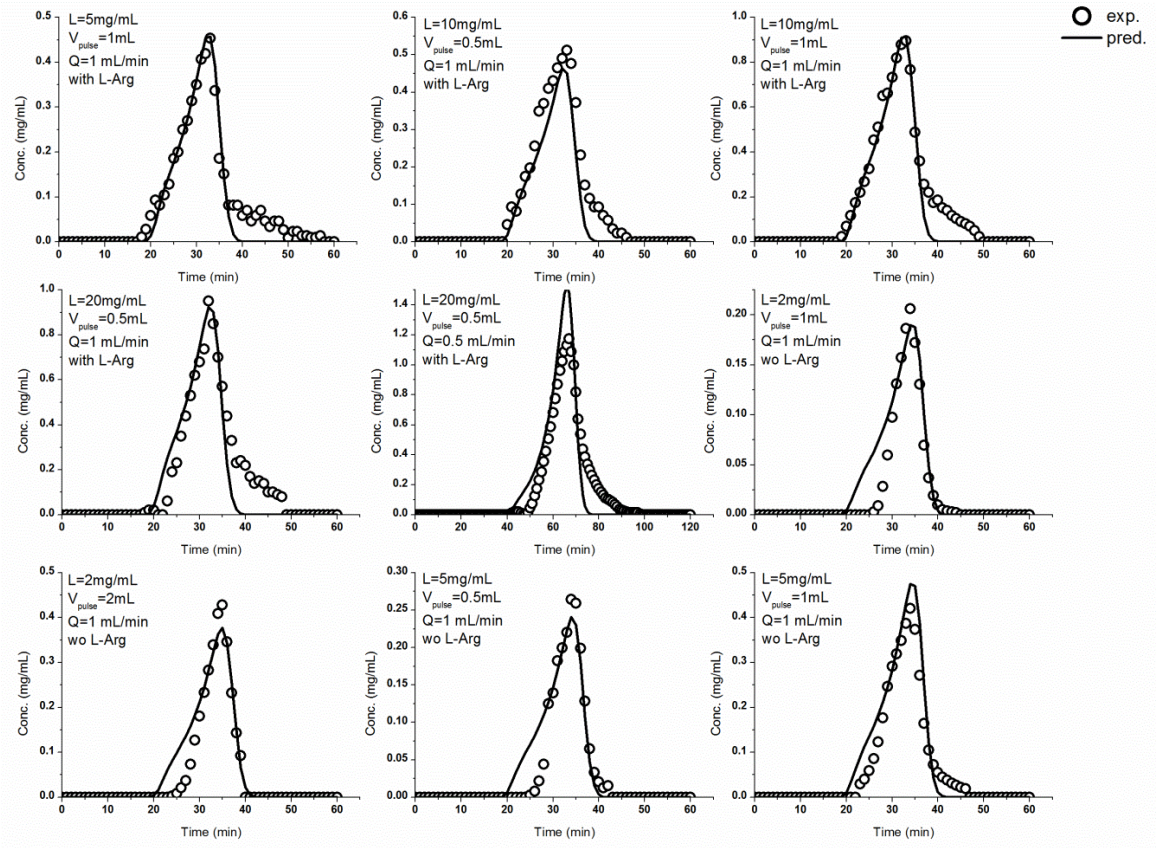


**Figure 3-8.** Experimental vs. predicted native protein elution profiles recovered during reactive experiments where an apparent two- state representation of reaction mechanism was used.

The apparent reaction rate constant was found to be 0.08±0.01 $min^{-1}$ and 0.1±0.01 $min^{-1}$ for refolding with buffer A and B respectively, indicating that the low concentration of L-arginine utilized in this work prevented aggregation without compromising the speed of the reaction (values are identical at 95% statistical significance). It should be noted that the disagreement between experimental and model prediction at the tail of the peak is due to non-specific interactions with the gel, which were not accounted for in the model. From these results it is evident that disappearance of the early intermediates is not immediate. Gradual removal of denaturing and reducing agents in SEC and the chemical composition used in this work may explain this behavior of the protein. For example, higher local concentration of chaotropic molecules such as urea has shown reduced apparent kinetics [33]. However, higher local concentration of such molecules and their

gradual removal protects early intermediate species against aggregation resulting in advantage of SEC over dilution refolding. In addition, carryover of excess concentration of reducing agent used in feed preparation can lead to suboptimum concentration ratio of redox couple and slow reformation of native disulfide bridges [8].

One should bear in mind that although the assumption of constant kinetics was proved applicable in this work, it may not be valid in higher loading volumes and reaction rate term may be more accurately determined via local concentrations of denaturing and reducing agents [6,33]. For example the full potential of SEC for protein refolding is realized in the application of continuous processing configurations such as simulated moving bed which normally results in less dilution and higher local concentrations [6]. Similarly, although no aggregates were formed at concentrations studied in this work the possibility of aggregation must be considered at higher loading volumes and local concentrations of protein.

## 3.5 Conclusion

The fundamental studies of oxidative protein refolding in SEC in this work showed: 1) an apparent two-state refolding mechanism adequately describes the refolding of model protein lysozyme on SEC highlighting the importance of chemical composition of the refolding environment and the rate by which this composition changes on kinetics of refolding; 2) different methods that can be used to find characteristic information about species involved in a refolding reaction such as quenched and equilibrium experiments and use of model compounds which is essential for model development and simulation of protein refolding on SEC; 3) examination of effect of L-arginine on characteristics of protein conformations showed higher apparent radius of the native lysozyme and no apparent change on larger protein conformations such as early intermediates and BSA at the presence of this additive; 4) low concentration of L-arginine used in this work (0.2 M) combined with 2 M urea showed insignificant effect on the rate of refolding reaction. This is of particular interest due to extensive application of these compounds as aggregation suppressor additives.

The same model can be extended to higher loading volumes or protein concentrations provided that the effect of local concentration of denaturing and reducing agents and local protein concentration on kinetics of refolding and aggregation are incorporated into the model.

## 3.6 References


[1] E.J. Freydell, L.A. van der Wielen, M.H. Eppink, M. Ottens, Techno-economic evaluation of an inclusion body solubilization and recombinant protein refolding process, Biotechnol. Prog. 27 (2011) 1315-1328.

[2] M. Wellhoefer, W. Sprinzl, R. Hahn, A. Jungbauer, Continuous processing of recombinant proteins: Integration of refolding and purification using simulated moving bed size-exclusion chromatography with buffer recycling, J. Chromatogr. A 1337 (2014) 48-56.

[3] M. Wellhoefer, W. Sprinzl, R. Hahn, A. Jungbauer, A. Continuous processing of recombinant proteins: Integration of inclusion body solubilization and refolding using simulated moving bed size exclusion chromatography with buffer recycling, J. Chromatogr. A 1319 (2013) 107-117.

[4] Y. Chen, S.S.J. Leong, High productivity refolding of an inclusion body protein using pulsed-fed size exclusion chromatography, Process Biochem. 45 (2010) 1570-1576.

[5] E.J. Freydell, L.A. van der Wielen, M.H. Eppink, M. Ottens, Size-exclusion chromatographic protein refolding: fundamentals, modeling and operation,J. Chromatogr. A 1217 (2010) 7723-7737.

[6] E.J. Freydell, Y. Bulsink, S. van Hateren, L. van der Wielen, M. Eppink, M. Ottens, Size-exclusion simulated moving bed chromatographic protein refolding,Chem. Eng. Sci. 65 (2010) 4701-4713.

[7] B.J. Park, C.H. Lee, S. Mun, Y.M. Koo, Novel application of simulated moving bed chromatography to protein refolding, Process Biochem. 41 (2006) 1072-1082.

[8] H. Lanckriet, A.P. Middelberg, Continuous chromatographic protein refolding, J. Chromatogr. A 1022 (2004) 103-113.

[9] Z. Gu, Z. Su, J.C. Janson, Urea gradient size-exclusion chromatography enhanced the yield of lysozyme refolding, J. Chromatogr. A 918 (2001) 311-318.

[10] B. Zelic, B. Nesek, Mathematical modeling of size exclusion chromatography, Eng. Life Sci. 6 (2006) 163-169.

[11] Z. Li, Y. Gu, T. Gu, Mathematical modeling and scale-up of size-exclusion chromatography, Biochem. Eng. J. 2 (1998) 145-155.

[12] Y. Ding, Y.P. Chuan, L. He, A.P. Middelberg, Dispersive mixing and intraparticle partitioning of protein in size-exclusion chromatographic refolding, J. Chromatogr. A 1218 (2011) 8503-8510.

[13] B. van den Berg, E.W. Chung, C. V Robinson, P.L. Mateo, C.M. Dobson, The oxidative refolding of hen lysozyme and its catalysis by protein disulfide isomerase, EMBO. J. 18 (1999) 4794-4803.

[14] A.M. Buswell, A.P. Middelberg, A new kinetic scheme for lysozyme refolding and aggregation, Biotechnol. Bioeng. 83 (2003) 567-577.

[15] M.E. Goldberg, R. Rudolph, R. Jaenicke,A kinetic study of the competition between renaturation and aggregation during the refolding of denatured-reduced egg white lysozyme, Biochemistry 30 (1991) 2790-2797.

[16] V. Vagenende, A.X. Han, M. Mueller, B.L. Trout, Protein-associated cation clusters in aqueous arginine solutions and their effects on protein stability and size, ACS Chem. Biol. 8 (2013) 416-422.

[17] D. Shukla, B.L. Trout, Interaction of arginine with proteins and the mechanism by which it inhibits aggregation, J. Phys. Chem. B 114 (2010) 13426-13438.

[18] J. Chen, Y. Liu, X. Li, Y. Wang, H. Ding, G. Ma, Z. Su, Cooperative effects of urea and L-arginine on protein refolding, Protein Expr. Purif. 66 (2009) 82-90.

[19] K.R. Reddy, H. Lilie, R. Rudolph, C. Lange, L-Arginine increases the solubility of unfolded species of hen egg white lysozyme, Protein Sci. 14 (2005) 929-935.

[20] T. Arakawa, K. Tsumoto, K. Nagase, D. Ejima, The effects of arginine on protein binding and elution in hydrophobic interaction and ion-exchange chromatography, Protein Expr. Purif. 54 (2007) 110-116.

[21] D. Ejima, R. Yumioka, T. Arakawa, K. Tsumoto, Arginine as an effective additive in gel permeation chromatography, J Chromatogr. A. 1094 (2005) 49-55.

[22] Y. Zhang, S. Rohani, A.K. Ray, Numerical determination of competitive adsorption isotherm of mandelic acid enantiomers on cellulose-based chiral stationary phase, J. Chromatogr. A. 1202 (2008) 34-39.

[23] S.W. Englander, L. Mayne, M.M. Krishna, Protein folding and misfolding: mechanism and principles,Q. Rev. Biophys. 40 (2007) 287-326.

[24] G. Wildegger, T. Kiefhaber, Three-state model for lysozyme folding: triangular folding mechanism with an energetically trapped intermediate, J. Mol. Biol. 270 (1997) 294-304.

[25] T. Kiefhaber, Kinetic traps in lysozyme folding, Proc. Natl. Acad. Sci. U S A 92 (1995) 9029-9033.

[26] V.P. Saxena, D.B. Wetlaufer, Formation of three-dimensional structure in proteins. I. Rapid nonenzymic reactivation of reduced lysozyme, Biochemistry 9 (1970) 5015-5023.

[27] A.M. Buswell, A.P. Middelberg, Critical analysis of lysozyme refolding kinetics,Biotechnol. Prog. 18 (2002) 470-475.

[28] R. Silvers, F. Sziegat, H. Tachibana, S. Segawa, S. Whittaker, U.L. Gunther, F. Gabel, J.R. Huang, M. Blackledge, J. Wirmer-Bartoschek, H. Schwalbe,Modulation of structure and dynamics by disulfide bond formation in unfolded states, J. Am. Chem. Soc. 134 (2012) 6846-6854.

[29] B. Raman, T. Ramakrishna, C.M. Rao, Refolding of denatured and denatured/reduced lysozyme at high concentrations, J. Biol. Chem. 271 (1996) 17067-17072.

[30] A.V.S. A. V. Volynskaya, S. A. Murasheva, A. Yu. Skripkin, A lysozyme unfolding mechanism, Dokl. Phys. Chem. 430 (2010) 29-32.

[31] P. Saremirad, J. A. Wood, Y. Zhang, A.K.Ray, Multi-variable operational characteristic studies of on-column oxidative protein refolding at high loading concentrations, J. Chromatogr. A 1359 (2014) 70-75.

[32] B. Batas, H.R. Jones, J.B. Chaudhuri, Studies of the hydrodynamic volume changes that occur during refolding of lysozyme using size-exclusion chromatography, J. Chromatogr. A 766 (1997) 109-119.

[33] S. Pan, M. Zelger, R. Hahn, A. Jungbauer, Continuous protein refolding in a tubular reactor, Chem. Eng. Sci. 116 (2014) 763-772.

# Chapter 4

# 4 Oxidative Protein Refolding on Size Exclusion Chromatography: From Batch Single-column to Multi-column Counter-current Continuous Processing


## Abstract

Recently size exclusion chromatography (SEC) used in multi-column continuous simulated moving bed (SMB) configurations (hereinafter SMB-SEC) has been investigated for protein refolding at industrial scale. This is due to several advantages offered by SMB configurations particularly when process parameters are thoroughly screened and optimized. A robust mathematical model is essential for high-throughput process screening and optimization. In this work, a previously investigated single-column mathematical model was modified to extend its applicability for protein oxidative refolding/aggregation predictions in SMB-SEC. The model considers a wider loading concentration range of the model protein (lysozyme) on SEC. The potential influences of high concentrations of chaotropic reagents on kinetic and thermodynamic model parameters have been discussed based on previous experimental results and their predicted local concentrations through the SMB-SEC columns and at the product steam. It was observed that aggregation occurs when local protein concentration exceeds a critical concentration. No urea recovery at the product stream indicated that the refolding reaction will continue off-column to recover the native- protein product. Therefore, it is suggested that the developed model is tested against experimental results for total solubilized protein (early intermediates and native conformations) and process performance indicators are defined based on solubilized protein.


## 4.1 Introduction

Despite the advances made to date for expression of protein-therapeutics using *E-coli* [1], existing technologies to recover active and high-purity product still incur significant costs due to low product concentration and high buffer consumption during conventional batch dilution refolding process resulting in low volumetric productivity [2]. As an alternative

due to the gradual separation of unfolded protein molecules from denaturing and reducing agents and their separation from solubilized aggregates in SEC, which in turn results in high refolding yield and purity, SEC has been widely used at lab scale [3–9]. However, this method in form of single-column batch processing may not enhance the aforementioned process performance indicators for industrial scale production of *E-Coli*-based protein-therapeutics. On the other hand, a multi-column continuous simulated moving bed system offers several advantages compared to single-column batch operation including increased productivity per unit mass of solid phase, lower solvent consumption, and less diluted products. This configuration consists of a set of chromatographic columns connected in series and is operated in continuous mode; the inlet/outlet lines are periodically shifted synchronously in the direction of liquid-phase flow to mimic countercurrent movement between a liquid-solvent and a solid phase. Multi-column continuous simulated moving bed is a well-established process for intensified difficult separations of small molecules and fine products e.g. separation of enantiomers [10,11]. However, its application for protein refolding and separation has recently attracted the attention of researchers [3–5,7]. For example, Freydell et al. [5] have reported a 35 times increase in productivity and 1/10 solvent consumption when a SMB-SEC was used for refolding of a model fusion protein compared to single-column processing.

Since many parameters are involved in operating a SMB-SEC (i.e. internal and external flow rates, switching time and feed concentration) systematic optimization studies are required to exploit the full potential of this system. A mathematical model is a pre-requisite for screening of process parameters and optimization. Freydell et al. [5] used a previously investigated single-column model in order to predict refolding/aggregation of a fusion protein in a four-zone SMB-SEC configuration comprised of multiple columns connected in series. They observed considerable discrepancy between model predictions and experimental results. For instance, the refolding yield was over-predicted by a factor of three. As discussed in the same work, this disagreement can be related to lower dilution factor in SMB-SEC compared to a batch single-column refolding resulting in higher local concentration of chaotropic agents (urea and DTT). And, in order to improve the SMB-SEC model- predictions in terms of native protein recovery the influence of local concentration of urea and DTT on model parameters should be considered.

The effect of lower dilution factor is twofold, as in addition to higher local concentrations of chaotropic reagents it also results in higher local protein concentration compared to a batch single-column refolding. Higher local protein concentration may additionally result in different reaction schemes. For example if for a single-column refolding experiment no aggregation was observed there would still remain the possibility of aggregation in SMB-SEC at the same protein loading concentrations.

In this work, (1) a previously experimentally verified single-column model was modified to expand its applicability for prediction of oxidative protein refolding/aggregation on SMB-SEC by considering a wider protein loading concentration range and the additional model parameters resulting from this modification were determined experimentally; (2) the sensitivity of refolding kinetics and possible complexity arising from reducing agent (DTT) carry-over have been discussed and DTT-free refolding was investigated and compared to previous studies with DTT carry-over; (3) the denaturing reagent (urea) mass transfer parameters were measured experimentally and used to predict the concentration of urea through SMB-SEC columns and at the product outlet under the current operation conditions; (4) the suitability of the developed model for process optimization was investigated; and (5) the effect of SMB-SEC operating parameters namely loading concentration and switching time on process performance indicators were predicted and the results were compared to single-column oxidative refolding of lysozyme.

## 4.2 Mathematical Model and Theory

### 4.2.1 Column Model

The protein refolding in size exclusion column was modeled using dispersive transport in the bulk with a film linear mass transfer resistance between particle-solid and bulk-liquid phases. The formulated differential mass balances for solutes in the bulk and the solid phases are [6].

$$\frac{\partial C_{b,i}}{\partial t} = D_L \frac{\partial^2 C_{b,i}}{\partial x^2} - u \frac{\partial C_{b,i}}{\partial x} - Pk_{ov,i}\left(C_{eqS,i} - C_{s,i}\right) + r_{b,i} \quad (4\text{-}1)$$

$$\frac{\partial C_{s,i}}{\partial t} = k_{ov}\left(C_{eqS,i} - C_{s,i}\right) + r_{s,i} \tag{4-2}$$

where $C_{b,i}$ and $C_{s,i}$ are the concentration of solute $i$ (unfolded, intermediates and native conformations) in bulk and solid phase respectively. In equations (4-1) and (4-2), $t$ represents time, $x$ axial distance along the column, $D_L$ axial dispersion coefficient, $u$ interstitial velocity, $k_{ov,i}$ solute overall mass transfer coefficient, $P$ phase ratio, $C_{eqS,i}$ the solid phase concentration in equilibrium with the bulk concentration. $r_{b,i}$ and $r_{s,i}$ are the net concentration change due to refolding and aggregation reactions in bulk and solid phases, which are described further in section 4.2.3.

The solute solid phase concentration in equilibrium with the bulk concentration was treated as a linear equilibrium relationship with a fixed equilibrium constant [12]:

$$C_{eqS,i} = K_{eq,i}C_{b,i} \tag{4-3}$$

where $K_{eq,i}$ is the equilibrium constant.

The boundary and initial conditions used to solve equations (4-1) and (4-2) are:

$$C_{b,i}(t, 0^-) = \begin{cases} C_{f,i} & 0 < t < t_{pulse} \\ 0 & t > t_{pulse} \end{cases} \tag{4-4a}$$

$$\frac{\partial C_{b,i}}{\partial x}(t, L_C) = 0 \tag{4-4b}$$

$$C_{b,i}(0, x) = 0 \tag{4-4c}$$

$$C_{s,i}(0, x) = 0 \tag{4-4d}$$

where $C_{f,i}$ is solute concentration in feed, $t_{pulse}$ is the duration of sample injection, and $L_C$ is the column length. The assumption that the sample is introduced into the column as a rectangular pulse of length $t_{pulse}$ was initially used as it has proven applicable in some cases [6,13]. However, the experimental injection profile was later introduced by a

Gaussian distribution function to further improve the accuracy of the model parameters and prediction results.

To solve the $C_{b,i}$ and $C_{s,i}$ along $x$ at different times, the first and second spatial derivatives were discretized as fourth-order central finite difference equations except for boundary points where second order forward and backward finite difference approximations were used. The resulting system of ODEs in time (method of lines) was solved numerically in MATLAB.

### 4.2.2 SMB-SEC Model

The selected design parameters of SBM-SEC model (e.g. number of columns in each zone and column dimensions) in this work were identical to the system used by Freydell et al. [5]. Figure 4-1 is a schematic representation of their system. Each zone comprises of two columns connected in series and as shown on the same figure an open loop system was used.

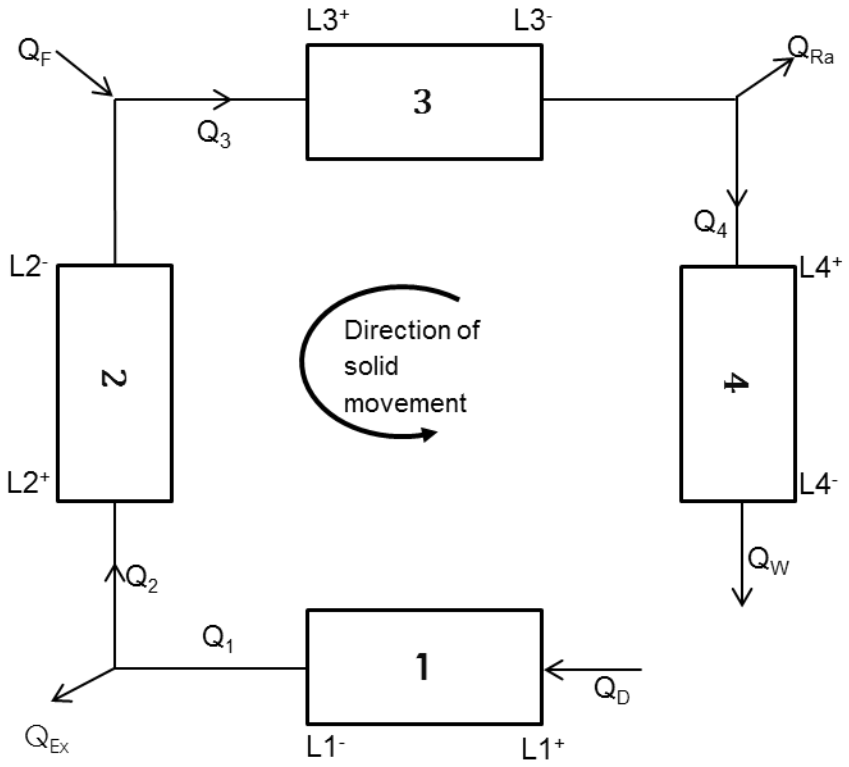


**Figure 4-1.** Schematic representation of a four-zone SMB-SEC: $Q_D$, $Q_{Ex}$, $Q_{Ra}$, $Q_F$ and $Q_W$ are buffer, extract, raffinate, feed and waste flow rates respectively. $Q_j$ is internal flow rate in zone j: 1-4)

The dimension of each column is 1 cm i.d and packed bed height of ~8 cm. Each column was modeled with the same approach as for single column except that boundary conditions are taken as periodic. The changing boundary condition is simulated by hypothetical movement of solid (direction shown on Figure 4-1) or more accurately switching columns' positions after each switching time. The boundary condition for each node is then:

$$C_{b,i}\,(t, L_1^+) = C_{D,i} \tag{4-5a}$$

$$C_{b,i}\,(t, L_2^+) = C_{b,i}\,(t, L_1^-) \tag{4-5b}$$

$$C_{b,i}\,(t, L_3^+) = \frac{Q_F C_{f,i} + Q_2 C_{b,i}\,(t, L_2^-)}{Q_3} \tag{4-5c}$$

$$C_{b,i}\,(t, L_4^+) = C_{b,i}\,(t, L_3^-) \tag{4-5d}$$

where $C_{D,i}$ is the concentration of solute $i$ in the refolding buffer.

### 4.2.3 Reaction Scheme

The refolding of reduced/denatured lysozyme has been successfully modeled using a single dominant reaction pathway in batch and fed batch refolding [14,15]. The reaction mechanism is shown in Figure 4-2.

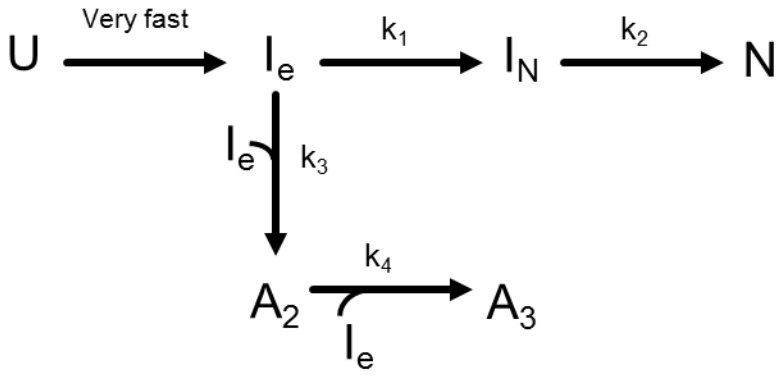


**Figure 4-2.** Lysozyme refolding and aggregation kinetic scheme, U: unfolded protein $I_e$: early intermediates, $I_N$ :native-like intermediates, N: native protein, $A_n$ : aggregates (n: 2 and 3).

It is suggested that early intermediates with only a small fraction of disulfide bridges are rapidly formed during "collapse process", followed by formation of native-like intermediates and a slow conversion of these intermediates to native state [16]. Since early intermediates are more susceptible to aggregation compared to native-like intermediates, aggregation is considered the result of early intermediate species association via sequential polymerization mechanism [6,14,15]. In our previous work, lysozyme refolding on SEC could be well described by an apparent two state mechanism involving early intermediates and native lysozyme [17]. Furthermore when lysozyme was

refolded without L-arginine additive, only one size of solubilized aggregates was recovered [18], therefore it was assumed that only one size of aggregate is also formed in the presence of L -arginine and aggregation was then regarded as a second order reaction. According to this simplified scheme the reaction rates for bulk ($r_{b,i}$) and solid ($r_{s,i}$) phases in equations (4-1) and (4-2) are as following

$$r_{I_e} = -k_{app} C_{I_e} - 2\, k_a (C_{I_e})^2 \tag{4-6a}$$

$$r_{I_N} = k_{app} C_{I_e} \tag{4-6b}$$

$$r_{A_2} = k_a (C_{I_e})^2 \tag{4-6c}$$

where $k_{app}$ and $k_a$ are apparent refolding and aggregation kinetic constants respectively.

Aggregation occurs at higher local concentrations which might be purely due to increased aggregation kinetics competing with refolding reaction or surpassing early refolding kinetic species solubility [19]. In the former case, the thermodynamic condition for aggregation formation is always satisfied while for the latter there is a critical concentration above which aggregates are formed. Elution profiles of refolded native protein for lysozyme refolding on SEC for a wide protein loading concentration range (5-50mg/mL) are examined in this work to investigate both case scenarios, assessed through inclusion of these effects in the model through an aggregation kinetic constant ($k_a$) and early intermediate solubility ($C_s$). If a critical concentration is required to describe the results of refolding/aggregating on SEC, the aggregation rate is modelled as following

$$r_{A_2} = \begin{cases} k_a (C_{I_e})^2 & C_{I_e} \geq C_s \\ 0 & C_{I_e} < C_s \end{cases} \tag{4-6d}$$

### 4.2.4 Model parameters estimated by single-column experiments

In our previous work an axial dispersion coefficient was calculated using available correlations for packed bed column. Also, the void volume, which was necessary for axial dispersion coefficient and phase ratio calculations, was measured using

thyroglobulin from bovine thyroid as a test probe. Furthermore, the mass transfer and equilibrium constants for early intermediates and native lysozyme ($k_{ov,i}$ and $K_{eq,i}$) were found by least-squares fitting of the deviation of measured concentration during single-column experiments vs. calculated for BSA model protein and native lysozyme respectively. The refolding kinetic constant ($k_{app}$) was also obtained in the same manner for elution profile of refolded native lysozyme when low loading concentrations (5-20 mg/mL) of denatured/reduced lysozyme were refolded on SEC. Low loading concentrations were used to prevent aggregation [17].

In this work, In order to introduce aggregation into the model, the aggregation kinetic constant ($k_a$) and early intermediate solubility ($C_s$), were determined by least squares fitting of measured refolded native protein concentration vs. calculated (equation 4-7) accomplished via the fminsearch function of MATLAB to find the parameters which indicated a global minimum.

$$f(x) = \sum_{j=1}^{j=n} (C_{b,i}^{exp,j} - C_{b,i}^{fit,j}(x))^2 \quad (4\text{-}7)$$

The above kinetic and thermodynamic parameters vary depending on local concentration of urea and DTT. In order to investigate the concentration profile of urea in SMB-SEC, mass transfer and equilibrium constants were found by measuring the concentration profile of pure urea injection on single column and minimization of deviation of experimental vs predicted results. The correlations for these constants have found to have wide errors as shown in the work of Park et al. [7], which motivated experimental determination of these values.

#### 4.2.4.1 Effect of Denaturant and Reducing Agents on Model Parameters

The carry-over of both urea and DTT, present in the feed stream, to the refolding environment can influence the results of refolding. However, it should be noted that DTT exists in both oxidized and reduced forms, which each possess different effects on refolding and aggregation kinetics. Concentrations equal or higher than of 0.6 mM of reduced form of this reagent can completely stop the refolding of lysozyme, while both the reduced and oxidized forms might result in non-optimal redox couple concentration

ratio and indirect adverse effects [8]. Furthermore, the oxidized to reduced ratio of this reducing agent will change over time due to oxidation unless continuous feed preparation is implemented. The oxidized to reduced ratio is also dependent on the feed protein concentration as the protein concentration varies whereas the initial concentration of DTT in the feed is constant. For the above mentioned reasons the task of quantifying the different effects becomes challenging. DTT may be removed before introducing the feed stream to the continuous refolding system. In order to gain more information on possible effects of DTT removal, batch single-column experiments of DTT-free lysozyme refolding were executed and the results were compared to the cases with carry-over of DTT.

The concentration profile of urea in SMB-SEC was predicted to find out more information about its local concentrations through the columns and separation from unfolded lysozyme. Since the purpose was to study the urea and protein concentration-waves separation, no reaction was considered. This investigation assisted in deciding whether further experimentations are required to find suitable functions to related kinetics of on-column refolding, aggregation and solubility to urea local concentrations.

## 4.3 Materials and Methods

### 4.3.1 Chemicals

Reagent grade L -arginine and urea, Ethylene Diamine Tetra Acetic acid (EDTA), lysozyme from chicken egg white, trizma® base (Tris-base), cysteine, cystine, BioUltradithiothreitol (DTT) solution, Trifluoric acetic acid (TFA) reagent plus grade and acetonitrile 0.1% TFA were purchased from Sigma-Aldrich, Canada. DIUR-500 urea assay kit and Red 660™ protein assay reagent were purchased from Bioassays, and G-Biosciences, USA respectively. Superdex™ 75pg resin (34 µm average particle size) was purchased from GE healthcare, Canada.

### 4.3.2 Analytical Methods

#### 4.3.2.1 Native-protein Concentration

A Vydac 214MS $C_4$ column (5 µm, 250 x 4.6 mm) was used on an Agilent HPLC system to separate native protein from other conformations and determine its concentration in collected samples during protein refolding experiments on SEC. A linear acetonitrile-water gradient with 0.1% (v/v) TFA starting at 25% acetonitrile increasing at 2.3%/min was used to elute the protein in 10 min. The total solvent flow rate, column temperature and injection volume were set at 1 mL/min, 20 ºC and 50 µL respectively.

Concentration of native protein in samples collected during injection profile determination was calculated by comparing the UV absorbance of samples and native protein standards at 280 nm. Both standards and samples UV absorbance reading were carried in a Tecan M200 plate reader.

#### 4.3.2.2 Total-protein Concentration

The total protein concentration in protein pool after DTT removal was determined with Red 660™ protein assay using microtiterplate reader in which 10 µl protein samples was transferred to each well, 150 µL of reagent was added and mixed using the plate reader (6.5 mm circular shaker), and absorbance of the mixture at 660 nm was measured after 5 min. The total protein concentrations were calculated by comparing the fraction and standard protein sample absorbance.

#### 4.3.2.3 Urea Concentration

The urea concentrations in samples were determined by urea assay kit (DIUR-500) using microtiterplate reader in which 5 µl of diluted samples were transferred to each well, 200 µL of reagent was added and mixed using the plate reader (6.5 mm circular shaker), and absorbance of the mixture at 520 nm was measured after 20 min. The samples were diluted with water instead of refolding buffer to avoid the interference of L -arginine with the assay (as recommended by the manufacturer).

### 4.3.3 Protein Unfolding

Unfolding buffer (0.1 M Tris-base, 1 mM EDTA, 6 M urea and 32 mM DTT, pH 8.1) was used to prepare various concentrations of denatured/reduced lysozyme. The sample was incubated for 2-4 h at 37 °C for lysozyme concentrations under 20 mg/mL [8] and above 30 mg/mL respectively. The loss of native structure was confirmed by RP-HPLC analysis afterwards.

### 4.3.4 Protein Injection Profile

The injection profile of protein was determined by replacing the column on ÄKTA purifier 100 by a piece of tubing with the same dimensions as column inlet tubing. The same injection loops were used for manual loading of native protein sample dissolved and eluted with refolding buffer. Fractions of 200 µl were collected and the concentration of native protein for each fraction was determined using UV absorbance.

### 4.3.5 DTT Removal

A 5 mL Hitrap desalting column (GE healthcare) was used to remove DTT from denatured/reduced lysozyme. The column was equilibrated with 2 CVs of unfolding buffer without DTT (0.1 M Tris-base, 1 mM EDTA, 6 M urea). 0.5 mL samples of 25 mg/mL were injected on column and eluted with 1 mL/min flow rate of unfolding buffer. Fractions of 0.5 mL were collected and pooled for refolding experiments described in the next section. In order to investigate the refolding of higher concentration of DTT-free lysozyme, the pool of desalting stage was concentrated using 10K 4 mL Amicon® Ultra centrifugal filters. The pool total protein concentration was determined using total protein assay as described above.

### 4.3.6 Single-column Batch Refolding

A XK16/40 column (GE healthcare, Canada) packed with Superdex™75pg resins (column volume ~54 mL) was installed on ÄKTA purifier 100 controlled by UNICORN 5.31 software and equipped with online pH probe, UV detector and conductivity cell. The

fractionation kit allows the collection of samples at desired volumes. Acetone pulse injection (2% (v/v)) was used to test the packing quality by comparing the peak symmetry and number of theoretical plates per length of column with manufacturer recommended criteria.

The column was equilibrated with 2 CVs of the refolding buffer (0.1 M Tris, 1 mM EDTA, 2 M urea, 0.2M L -arginine, 3 mM cysteine, 0.3 mM cysteine buffered at pH 8.1) prior to injection of 1 mL of denatured/reduced lysozyme with various concentrations (30, 40 and 50 mg/mL) . The sample was eluted using the same buffer and flow rate of 1 mL/min. Fractions of 1 mL were collected and concentrations of native protein were determined using RPHPLC. The same procedure was followed for refolding of denatured/reduced lysozyme after DTT removal.

### 4.3.7 Solubility Test

Samples of 40 mg/mL DTT-free lysozyme in 0.1 M Tris-base, 1 mM EDTA, 6 M urea buffered at pH 8.1 were diluted using a 0.1 M Tris-base, 1 mM EDTA buffer pH 8.1containing urea and L-arginine without redox couple. The urea and L-arginine concentrations were adjusted to result in a buffer identical to refolding buffer used during on-column refolding experiments considering the dilution factor. The absence of redox couple quenches the reaction at early intermediates [18]. DTT was removed for the same reason to avoid the formation of native protein during dilution experiments. The initial concentration of diluted sample was 8 mg/mL which was incubated overnight. The aggregates were precipitated by centrifuge and the concentration of solubilized protein in supernatant was measured by total protein assay.

### 4.3.8 Urea Injection

A sample of 6 M urea was prepared in a buffer similar to the refolding buffer composition with the exception of urea. No urea was used in this buffer to avoid high dilution factors during urea concentration measurements. 0.5 mL of this sample was injected on the SEC column and eluted with 1 mL/min flow rate of the same buffer while

collecting fractions of 1 mL. The experimental elution profile of urea was obtained by measuring urea concentration in the samples using urea assay kit.

### 4.3.9 SMB-SEC Simulations

After determining the necessary model parameters, the SMB-SEC operational parameters (i.e. external and internal flow rates and switching time) were selected using triangle theory. Based on this theory, inequalities of (4-8a)(a-c) correspond to a design space where complete separation can be achieved [20].

$$K_{eq,urea} < m_1 < \infty \tag{4-8a}$$

$$K_{eq,N} < m_2 < m_3 < K_{eq,urea} \tag{4-8b}$$

$$\frac{-\varepsilon_p}{(1-\varepsilon_p)} < m_4 < K_{eq,urea} \tag{4-8c}$$

where $m_j$ is called flow rate ratio and defined as ratio of net liquid flow rate to net solid phase flow rate in zone $j$ and $\varepsilon_p$ is particle porosity. Particle porosity was measured experimentally using acetone and thyroglobulin from bovine thyroid injections on single column. Acetone and thyroglobulin from bovine thyroid elution volume to total column volume ratio measures total porosity ($\varepsilon$) and void volume ($\varepsilon_b$) respectively and the particle porosity is calculated as: $\varepsilon_p = \frac{\varepsilon-\varepsilon_b}{1-\varepsilon_b}$. Total porosity, void volume and particle porosity were found to be 0.9, 0.34 and 0.8 respectively. The equilibrium constants for native lysozyme and urea were found by least-squares fitting of the deviation of measured concentration vs. calculated for native protein and urea when native lysozyme and pure urea were injected on the single column. The fitted parameters are 0.9 for urea and 0.44 for native lysozyme [17]. The liquid flow rate in each SMB-SEC zone is related to phase ratio by the following

$$Q_j = \frac{m_j V_c(1-\varepsilon) + V_c\varepsilon}{t_s} \tag{4-9}$$

where $t_s$ is the switching time.

In addition to the above criteria the flow rates must be within lowest and the highest flow rate range which are considered 0.25 and 3 mL/min respectively. For the selected flow rate range, Young and Wilson-Geankoplis correlations [6] were used to determine the controlling mass transfer resistance by order of magnitude analysis which revealed the pore resistance as the major resistance component. Since this value is independent of the velocity the estimated model parameters are valid for the selected flow rate range. The highest flow rate is determined based on maximum allowable pressure over the system compatible with packing and column material. Table 4-1 reports the operational parameters which were used during SMB-SEC simulations.

**Table 4-1.** SMB-SEC operational parameter selected based on triangle theory

| Phase ratios | $t_s$ (min) | $Q_D$=$Q_1$ (mL/min) | $Q_2$ (mL/min) | $Q_3$ (mL/min) | $Q_4$=$Q_W$ (mL/min) | $Q_{Ex}$ (mL/min) | $Q_F$ (mL/min) | $Q_{Ra}$ (mL/min) |
|---|---|---|---|---|---|---|---|---|
| $m_1$=1.3 | 2.7 | 2.72 | 1.66 | 1.96 | 1.21 | 1.04 | 0.30 | 0.74 |
| $m_2$=0.6 | 3 | 2.39 | 1.47 | 1.74 | 1.08 | 0.92 | 0.26 | 0.66 |
| $m_3$=0.8 $m_4$=0.3 | 3.3 | 2.17 | 1.34 | 1.58 | 0.98 | 0.84 | 0.24 | 0.64 |

The continuous and batch process performance indicators are defined as:

$$R_{b,exp} = \frac{M_{soluble}}{LV_{pulse}} \tag{4-10a}$$

$$R_{c,pred} = \frac{C^{CSS}_{soluble.Ra} Q_{Ra}}{Q_F C_{f,U}} \tag{4-10b}$$

Where $R_b$ and $R_{c,pred}$ are experimental and predicted solubilized lysozyme recovery for batch and continuous refolding respectively. $M_{soluble}$ is sum of solubilized protein mass collected in fractions, $V_{pulse}$ injection volume, $L$ lysozyme loading concentration and $C^{CSS}_{soluble.Ra}$ predicted average concentration of solubilized protein at cyclic steady state.

$$Pr_{b,exp} = \frac{M_{soluble}}{V_C t_e} \quad (4\text{-}11a)$$

$$Pr_{c,pred} = \frac{C^{CSS}_{soluble.Ra} Q_{Ra}}{N_t V_C} \quad (4\text{-}11b)$$

where $Pr_{b,exp}$, $Pr_{c,pred}$, $t_e$, $V_c$, and $N_t$ are experimental and predicted volumetric productivity for batch and continuous configurations, elution time of protein for batch experiments, volume of the column and total number of columns used for SMB-SEC respectively.

$$P_{c,pred} = \frac{C^{CSS}_{soluble.Ra}}{C^{CSS}_{soluble.Ra} + C^{CSS}_{urea.Ra}} \quad (4\text{-}12)$$

$P_{c,pred}$ is the predicted product purity.

$$Bc_{b,exp} = \frac{Q_b t_e}{M_{soluble}} \quad (4\text{-}13a)$$

$$Bc_{c,pred} = \frac{Q_D}{C^{CSS}_{Soluble.Ra} Q_{Ra}} \quad (4\text{-}13b)$$

where $Bc_{b,exp}$ and $Bc_{c,pred}$ are experimental and predicted buffer consumption and $Q_b$ is elution flow rate used for single column batch experiments.

## 4.4 Results and Discussions

### 4.4.1 Single-column Batch Refolding

When 1 mL of 30, 40 and 50 mg/mL of denatured/reduced lysozyme was refolded on SEC, although no protein signal was observed for aggregates, the refolding yield values (70, 50 and 40%) suggested formation of insoluble aggregates and in-column precipitation of these species. In order to decide whether aggregation is kinetically controlled by competition of refolding and aggregation or if a critical concentration is required, two models were examined to predict the results of lysozyme

refolding/aggregation on SEC for the concentration ranges used in this work as well as our previous work (described in section 4.2.3) [17] . Figure 4-3 shows that the assumption that thermodynamic condition for aggregation formation is always satisfied does not provide a satisfactory agreement between experimental and predicted results.

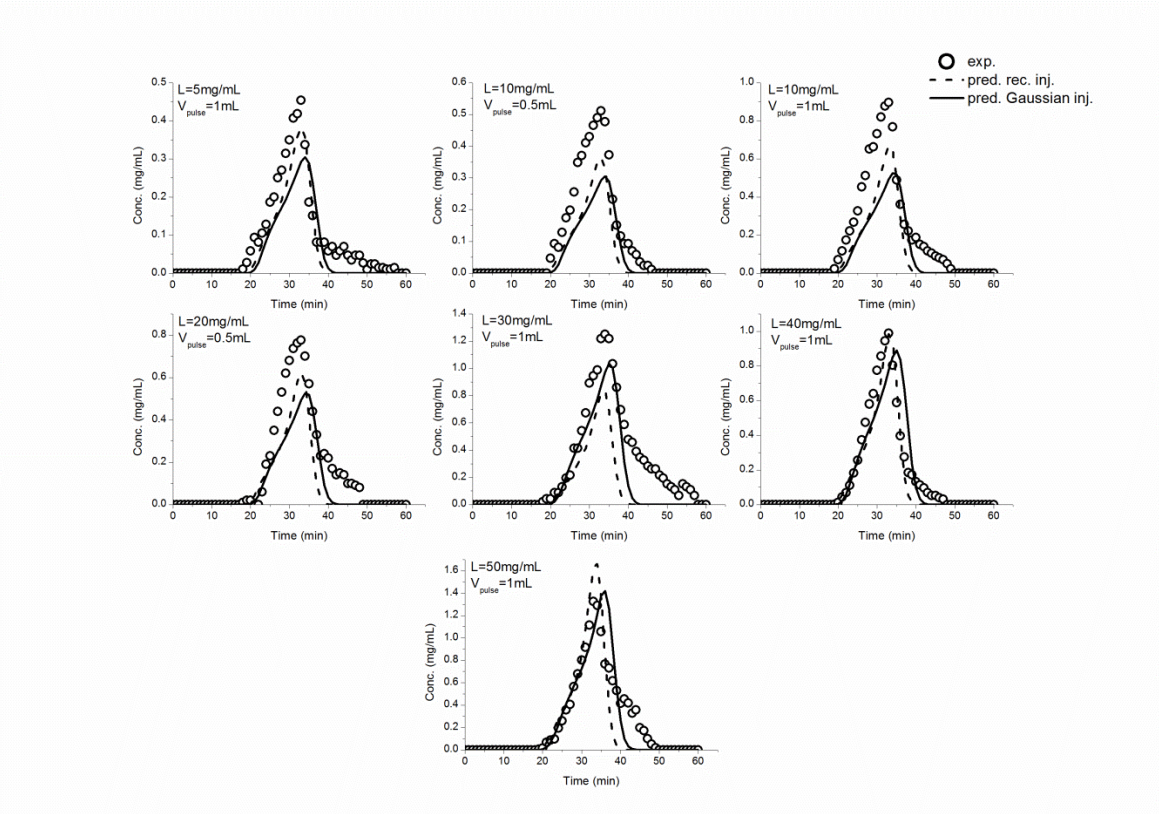


**Figure 4-3.** Experimental vs predicted elution profile of native refolded lysozyme assuming thermodynamic condition for aggregation formation is always satisfied**.**

Since the competing reaction rates are dependent on local concentration of protein the experimental injection profile of the protein was measured and used instead of commonly used simplified rectangular injection profile to provide a more accurate representation of the local protein concentrations. As shown in Figure 4-4 Gaussian type functions provide a good representation of the injection profiles of the protein in the system under study. It is evident from the same figure that the actual injection profile significantly deviates from the rectangular assumption. Using the experimental injection profile however, did not

change previously estimated model parameters namely mass transfer and equilibrium constants for kinetic species (data not shown).

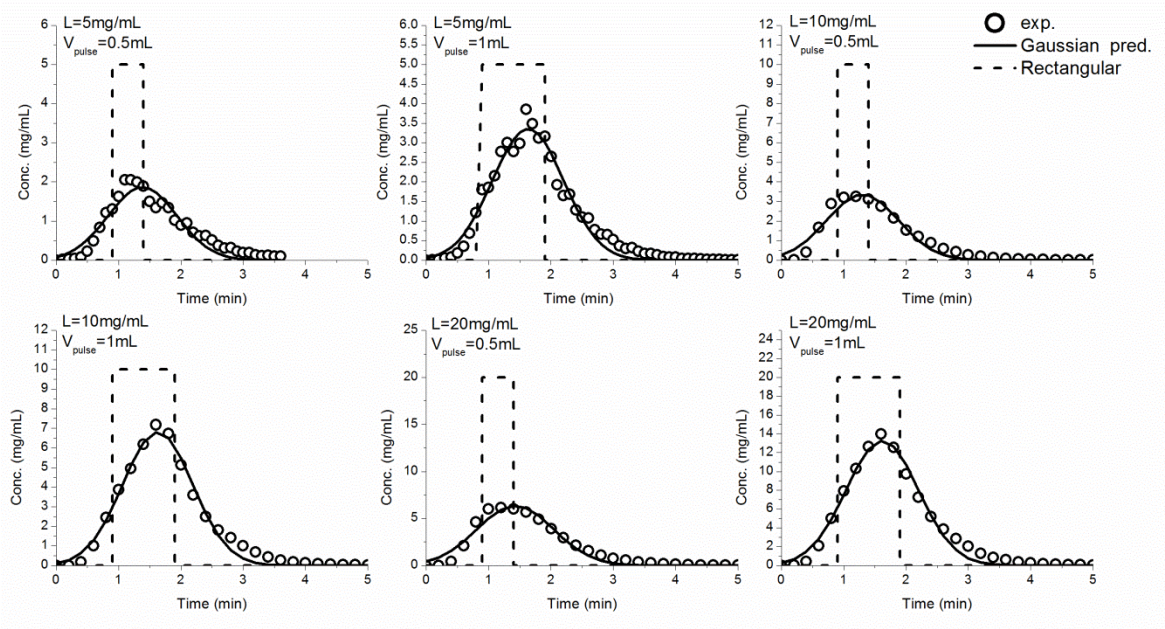


**Figure 4-4.** Experimental injection profile of lysozyme and its comparison with simplified rectangular injection profile.

The experimental injection profile was then introduced to the model. Nevertheless, the model prediction was not satisfactory (as shown on Figure 4-3) suggesting that the model is not adequate and thermodynamic condition for aggregation is not satisfied for the whole concentration range. In other words, aggregation should be only introduced to the model when the local protein concentration exceeds a critical concentration (i.e. solubility of early intermediate species). As illustrated in Figure 4-5 the modified model considering a critical concentration and experimental injection profile of the protein provides an improved agreement between experimental and predicted results. The two fitted parameters namely aggregation kinetic constant ($k_a$) and early intermediate solubility ($C_s$) however were found to be correlated and multiple solutions were obtained for minimization problem. Consequently, the solubility parameter was measured experimentally as described in section 4.3.7. The experimental and fitted parameters were 4.4±0.9 mg/mL and 0.05±2.8 x $10^{-5}$ mL/mg min respectively.

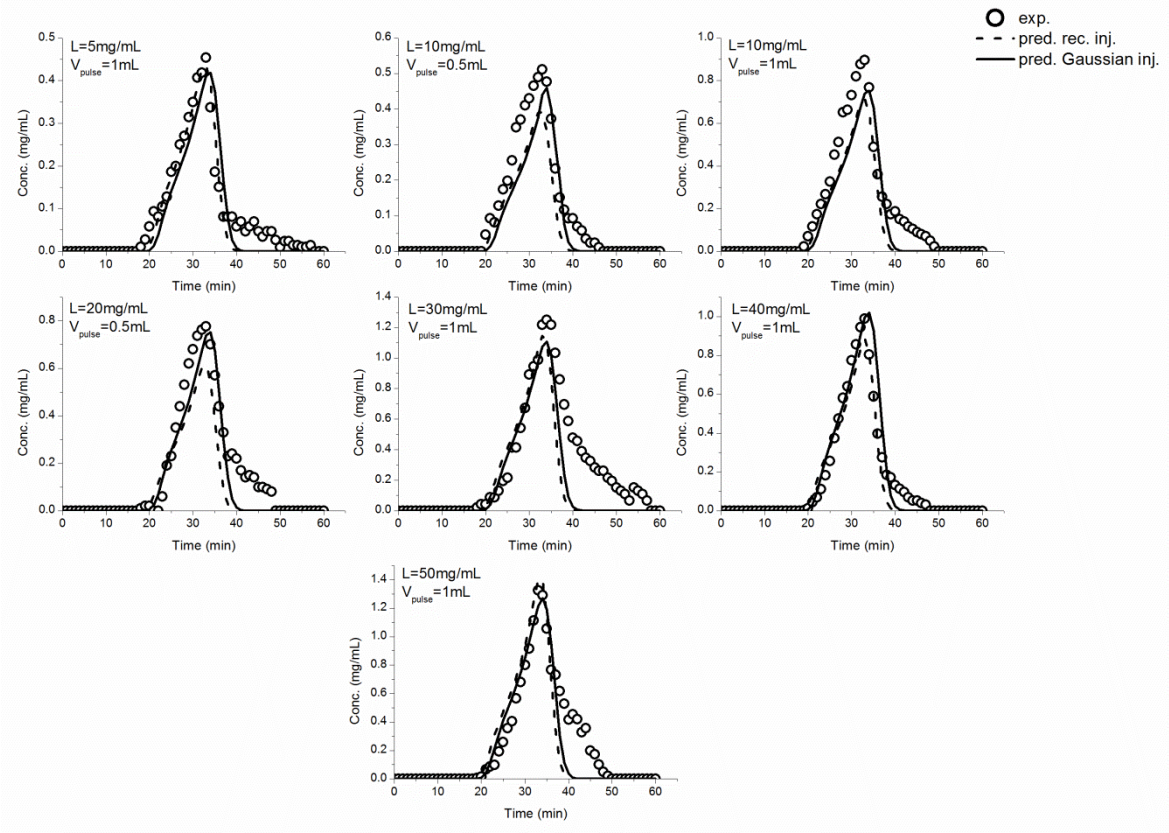


**Figure 4-5.** Experimental vs predicted elution profile of native refolded lysozyme using both simplified rectangular and experimental injection profiles.

### 4.4.2 Single-column Batch Refolding of DTT-free Lysozyme

The apparent refolding kinetic of DTT-free lysozyme found to be equal to lysozyme refolding with carry-over of DTT within experimental error. It can be seen in Figure 4-6 that $k_{app}$=0.08 min$^{-1}$ adequately predicts the result of DTT-free lysozyme refolding on SEC at loading concentrations of ~5 and 10 mg/mL and 1 mL injection volume. This is because the refolding kinetic constant with DTT carry-over was measured for a system with high dilution factor (~50 times) and the carry-over of reduced DTT was insignificant.

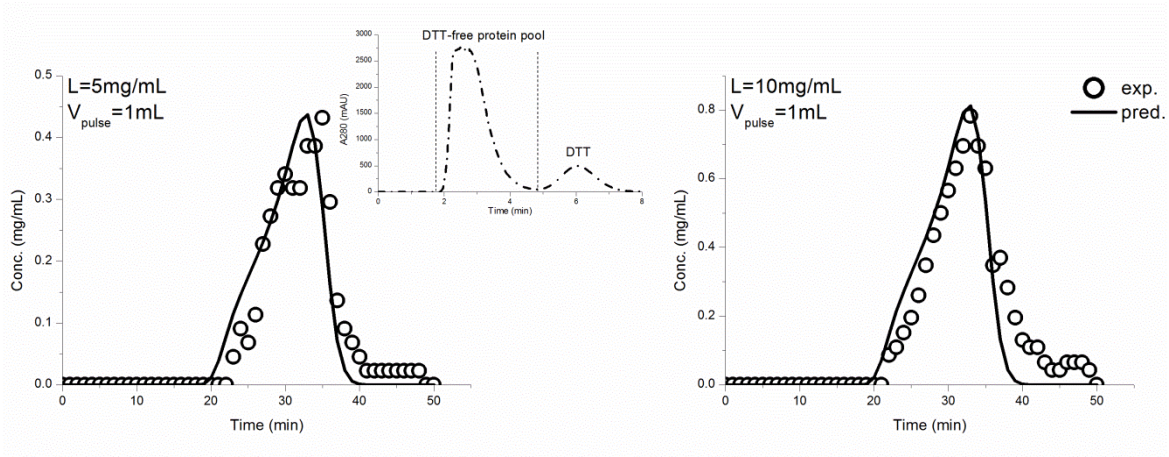


**Figure 4-6.** Experimental vs predicted elution profile of refolded native lysozyme for DTT- free lysozyme loading concentrations of 5 mg/mL and 10 mg/mL; DTT-free lysozyme obtained from desalting column was pooled, concentrated and injected on SEC.

### 4.4.3 Model Parameters for Urea

The experimental mass transfer and equilibrium constants for urea ($k_{ov,urea}$ and $K_{eq,urea}$) were found to be 51.97±0.02 $min^{-1}$ and 0.9±0.1. Figure 4-7 illustrates the experimental versus predicted urea concentration at the column outlet using experimentally measured mass transfer parameters as well as parameters calculated by available correlations. It was observed that correlations do not provide an accurate prediction of urea elution profile. For this reason, the fitted values of mass transfer parameters were applied to predict the local concentrations of urea through the SMB-SEC columns and at the raffinate stream.

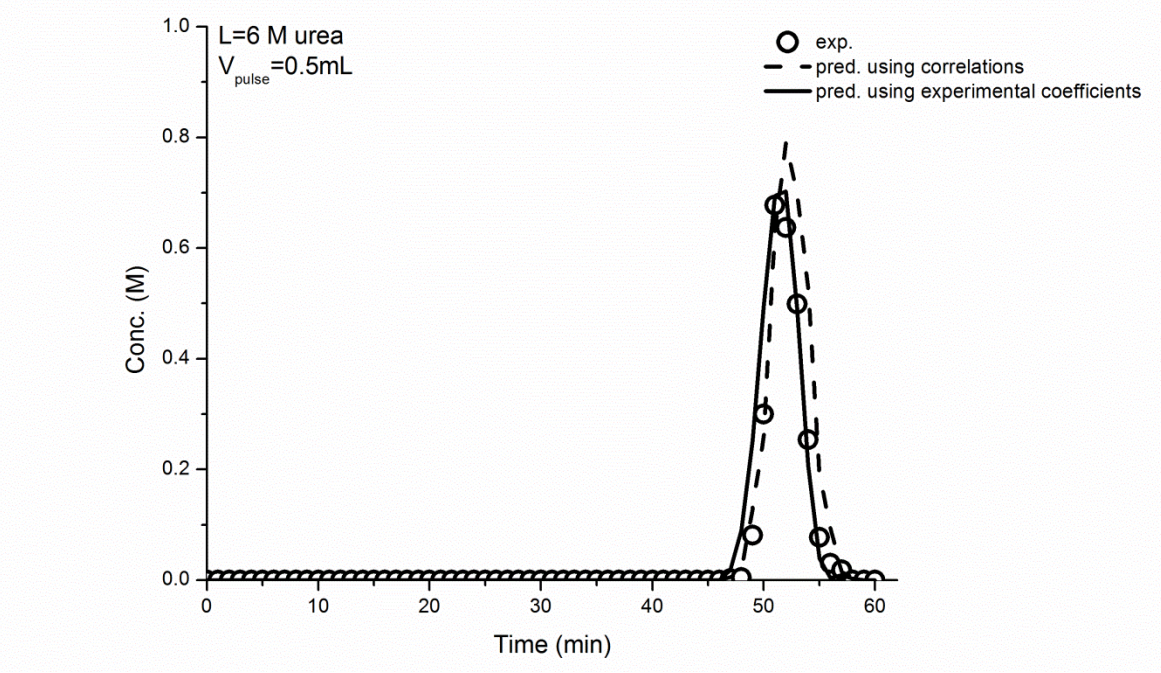


**Figure 4-7.** Experimental vs predicted elution profile of urea; predicted with the coefficients taken from correlations and coefficients taken from independent experiments.

### 4.4.4 SMB-SEC Model Validation

Park et al. [7] have studied the refolding of DTT-free lysozyme (1.78 mg/mL) in a 1-1-1-1 four zone SMB-SEC. The model developed in the current work, was initially tested against their experimental results and native protein concentration and refolding yield at product outlet was predicted with 9.3% and 5.5% relative error respectively. Although, the discrepancy between model prediction and existing experimental results is not significant, the model must be tested under various operation conditions such as loading concentrations for further validation.

It can be seen from Figure 4-8 that for the configuration used in this work, there is an overlap between urea and protein concentrations. The overlapping area corresponds to the length of two columns in SMB-SEC. This observation suggest that more experimentation may be required to relate the model parameters namely apparent refolding kinetic constant, solubility of early intermediates and aggregation constant to local concentration

of urea in order to account for the effect of dynamic chemical environment on these parameters [21].

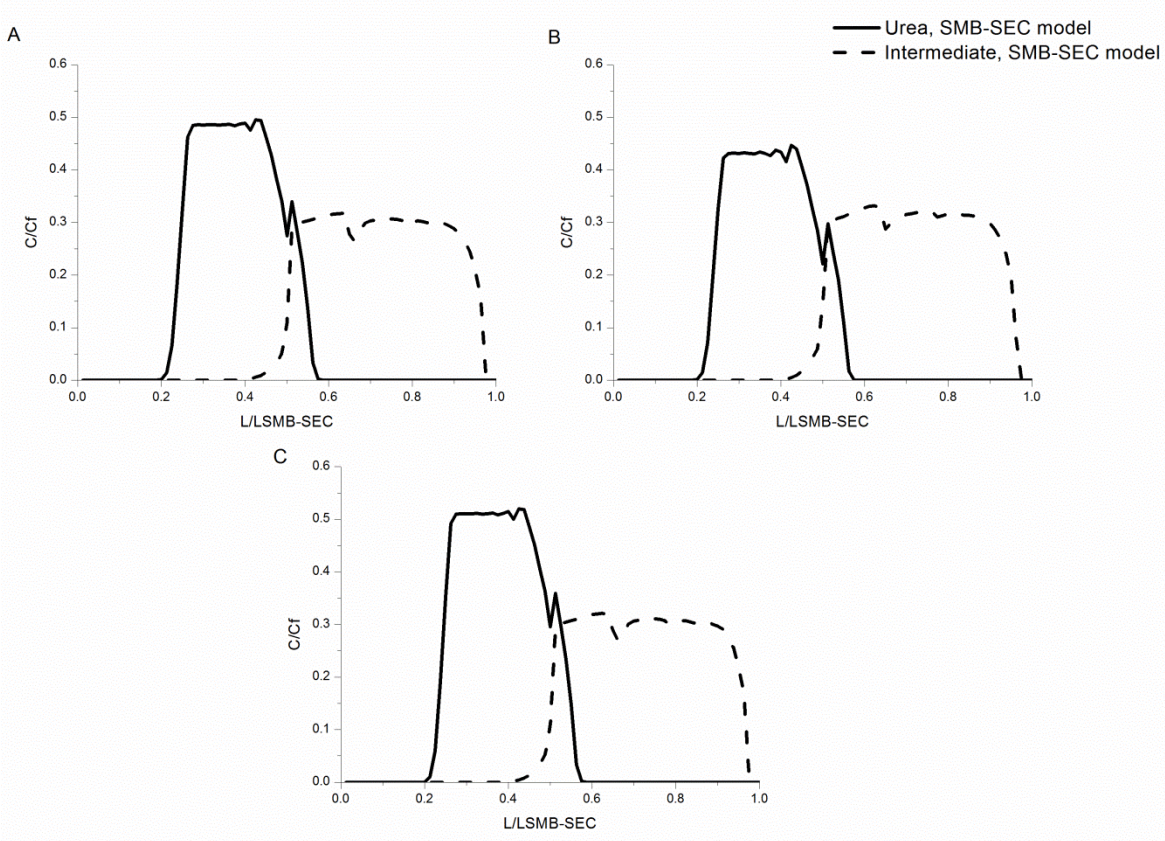


**Figure 4-8.** Mid cycle concentration profiles of urea and early intermediates through SMB-SEC columns for the operating condition reported in Table 4-1 and 10 mg/mL DTT-free lysozyme loading concentration (The number of switching time was fixed at 30 in order to assure steady state operation); (A) $t_s$ =2.7 min, (B) $t_s$ =3 min and (C) $t_s$ =3.3 min.

Urea has demonstrated both denaturing and protective osmolyte properties [22,23]. For example, no native conformation has been observed for lysozyme at urea concentrations above 5 M [24] whereas urea concentrations under 3 M seem to increase the stability of the collapsed-globular state of poly (N-isopropylacrylamide) and shift the unfolded/folded equilibrium towards folded conformations [22]. Under the current operating conditions, as can be seen in Figure 8, urea concentration varies along the columns between no excess urea compared to the refolding buffer (total urea

concentration of 2M) and about 50% of the urea present in the feed (total urea concentration of about 5M). Since the model parameters were measured using the refolding buffer during single-column experiments with much higher dilution factor, the developed model with constant kinetic and thermodynamic parameters may over-predict the native protein concentration at SMB-SEC product outlet as it disregards lower refolding kinetics and equilibrium in higher urea concentrations. Over-prediction of the native protein concentration results in under-prediction of protein aggregation. However, this error may be balanced by higher solubility of early intermediates and reduced aggregation in higher urea concentrations.

On the other hand, simulations for the above operating conditions showed no urea in the raffinate stream, indicating the refolding reaction will continue off-column with the same kinetics as single column experiments to recover the desired product. Therefore in this work, process performance indicators were defined based on total solubilized protein rather than native protein. And accordingly, it is suggested that the model is tested against experiments for total solubilized protein. It should be noted that in the presence of L-arginine, aggregates precipitate and solubilized protein only include early intermediates and native protein conformations.

### 4.4.5 SMB-SEC Performance

Based on above discussion, the model with constant parameters was used to study the effect of operating parameters on process performance indicators namely solubilized protein recovery and productivity, product purity at the raffinate outlet and buffer consumption . Table 4-2 shows the results of SMB-SEC simulations at operating condition corresponding to switching time of 3 min in Table 4-1 for feed concentrations of 5-40 mg/mL. The corresponding SEC experimental results are also reported when data was available from experiments.

**Table 4-2.** SMB-SEC simulations feed concentrations of 5-40 mg/mL, switching time of 3 min and available corresponding SEC experimental results. SEC results are reported for elution flow rate of 1 mL/min and injection volume of 0.5 mL.

| $C_{f,U}$ (mg/mL) | $R_{c,pred}$ (-) | $Pr_{c,pred}$ (mg/mL h) | $Bc_{c,pred}$ (mL/mg) | $R_{b,exp}$ (-) | $Pr_{b,exp}$ (mg/mL h) | $Bc_{b,exp}$ (mL/mg) |
|---|---|---|---|---|---|---|
| 5 | 1.0 | 1.6 | 1.9 | 1.0 | 0.1 | 13.6 |
| 10 | 1.0 | 3.0 | 1.0 | 1.0 | 0.2 | 6.8 |
| 15 | 1.0 | 4.8 | 0.6 | - | - | - |
| 20 | 0.8 | 5.4 | 0.5 | 1.0 | 0.3 | 3.4 |
| 40 | 0.5 | 6.0 | 0.5 | 0.9 | 0.6 | 1.9 |

It can be seen that, although productivity increases and buffer consumption is reduced at higher feed concentrations, the recovery decreases for concentrations equal and above 20 mg/mL showing aggregation formation. As expected the productivity and buffer consumption in SMB-SEC were significantly improved compared to SEC. The buffer consumption can be further reduced by using a closed loop configuration and buffer recycling. For oxidative refolding however, the redox couple ratio might change over time resulting in non-optimal buffer composition. It was previously demonstrated that lysozyme refolds at higher pH values closer to its isoelectric point without the need for redox couple while aggregation was suppressed due to the presence of L -arginine and the refolding yield did not show a significant change [18]. Based on the characteristics of the protein under study oxidative refolding without the need for a redox couple might offer advantages in terms of chemical cost, buffer preparation and storage but further kinetic studies are required to compare refolding with and without a redox couple. As mentioned 100% purity was estimated for SMB-SEC as predictions showed no urea at the raffinate stream.

In order to explore the effect of switching time, the simulations were also carried out for the above feed concentration range and operating conditions corresponding to switching times of 2.7 and 3.3 min. As shown in Table 4-3, in the studied concentration range switching time of 2.7 min resulted in highest productivity compared to switching times of 3 and 3.3 min and up to 50% improvement was observed for loading concentration of 5 mg/mL, while recovery and buffer consumption did not vary significantly by change of the switching time. The same trend was observed for switching times of 2.7 and 3.3 min in terms of recovery drop for concentrations equal and above 20 mg/mL as switching time of 3 min.

**Table 4-3.** SMB-SEC productivity simulations for feed concentrations of 5-40 mg/mL and various operation conditions corresponding to three different switching times used in this work.

| $C_{f,U}$ (mg/mL) | $t_s$ (min) | $Pr_{c,pred}$ (mg/mL h) |
|---|---|---|
| 5 | 2.7 | 1.8 |
| 10 | 2.7 | 3.6 |
| 15 | 2.7 | 5.4 |
| 20 | 2.7 | 6.0 |
| 40 | 2.7 | 7.2 |
| 5 | 3 | 1.6 |
| 10 | 3 | 3.0 |
| 15 | 3 | 4.8 |
| 20 | 3 | 5.4 |
| 40 | 3 | 6.0 |
| 5 | 3.3 | 1.2 |
| 10 | 3.3 | 3.0 |
| 15 | 3.3 | 4.2 |
| 20 | 3.3 | 5.4 |
| 40 | 3.3 | 6.0 |

## 4.5 Conclusions

This work illustrates important considerations for utilizing single-column data towards design/operation of an SMB process. Our findings showed: 1) at higher local protein concentrations aggregation occurs when local protein concentration exceeds a critical concentration (i.e. solubility of early intermediates); 2) if DTT is to be removed from denatured/reduced lysozyme to avoid further complexity in developing a model and the adverse effect of this reagent on refolding kinetics, the single-column parameters obtained with DTT carry-over are still valid provided they were obtained for a column with high dilution factor; 3) a model with constant kinetic and thermodynamic parameters may not result in accurate prediction of native protein; however, 4) under the operation condition studied, the refolding reaction will continue off-column as predictions showed no urea in the raffinate stream ; and 5) it is suggested that the model is tested against total solubilized protein rather than native protein and the process performance indicators are defined accordingly; finally 6) the prediction of SMB-SEC performance for loading concentrations of 5-40 mg/mL, using the model with constant parameters, showed that increasing the concentration increases the productivity and decreases the buffer consumption however at concentrations equal to and higher than 20 mg/mL aggregates are formed and precipitate in the column. The operating condition corresponding to lowest switching time resulted in highest productivity and no significant effect on recovery and buffer consumption.

## 4.6 References


[1] C.J. Huang, H. Lin, X. Yang, Industrial production of recombinant therapeutics in Escherichia coli and its recent advancements, J. Ind. Microbiol. Biotechnol. 39 (2012) 383-399.

[2] E.J. Freydell, L.A. van der Wielen, M.H. Eppink, M. Ottens, Techno-economic evaluation of an inclusion body solubilization and recombinant protein refolding process, Biotechnol. Prog. 27 (2011) 1315-1328.

[3] M. Wellhoefer, W. Sprinzl, R. Hahn, A. Jungbauer, Continuous processing of recombinant proteins: Integration of refolding and purification using simulated

moving bed size-exclusion chromatography with buffer recycling , J. Chromatogr. A 1337 (2014) 48-56.

[4] M. Wellhoefer, W. Sprinzl, R. Hahn, A. Jungbauer, Continuous processing of recombinant proteins: Integration of inclusion body solubilization and refolding using simulated moving bed size exclusion chromatography with buffer recycling, J. Chromatogr. A 1319 (2013) 107-117.

[5] E.J. Freydell, Y. Bulsink, S. van Hateren, L. van der Wielen, M. Eppink, M. Ottens, Size-exclusion simulated moving bed chromatographic protein refolding, Chem. Eng. Sci. 65 (2010) 4701-4713.

[6] E.J. Freydell, L.A. van der Wielen, M.H. Eppink, M. Ottens, Size-exclusion chromatographic protein refolding: fundamentals, modeling and operation , J Chromatogr A 1217 (2010) 7723-7737.

[7] B.J. Park, C.H. Lee, S. Mun, Y.M. Koo, Novel application of simulated moving bed chromatography to protein refolding , Process Biochem. 41 (2006) 1072-1082.

[8] H. Lanckriet, A.P. Middelberg, Continuous chromatographic protein refolding , J. Chromatogr. A 1022 (2004) 103-113.

[9] Z. Gu, Z. Su, J.C. Janson, Urea gradient size-exclusion chromatography enhanced the yield of lysozyme refolding, J. Chromatogr. A 918 (2001) 311-318.

[10] A. Rajendran, G. Paredes, M. Mazzotti, Simulated moving bed chromatography for the separation of enantiomers, J. Chromatogr. A 1216 (2009) 709-738.

[11] G. Ströhlein, M. Mazzotti, M. Morbidelli, Simulated Moving-Bed Reactors, in:, Integr. Chem. Process. Synth. Oper. Anal. Control, Wiley-VCH Verlag GmbH & Co. KGaA, 2005, pp. 183–201.

[12] B. Zelic, B. Nesek, Mathematical modeling of size exclusion chromatography, Eng. Life Sci. 6 (2006) 163-169.

[13] Y. Zhang, S. Rohani, A.K. Ray, Numerical determination of competitive adsorption isotherm of mandelic acid enantiomers on cellulose-based chiral stationary phase, J. Chromatogr. A 1202 (2008) 34-39.

[14] A.M. Buswell, A.P. Middelberg, A new kinetic scheme for lysozyme refolding and aggregation, Biotechnol. Bioeng. 83 (2003) 567-577.

[15] M.E. Goldberg, R. Rudolph, R. Jaenicke, A kinetic study of the competition between renaturation and aggregation during the refolding of denatured-reduced egg white lysozyme, Biochemistry 30 (1991) 2790-2797.

[16] B. van den Berg, E.W. Chung, C. V Robinson, P.L. Mateo, C.M. Dobson, The oxidative refolding of hen lysozyme and its catalysis by protein disulfide isomerase, EMBO J 18 (1999) 4794-4803.

[17] P. Saremirad, J.A. Wood, Y. Zhang, A.K. Ray, Oxidative Protein Refolding on Size Exclusion Chromatography at High Loading Concentrations: Fundamental Studies and Mathematical Modeling, J. Chromatogr. A 1370 (2014) 147-155.

[18] P. Saremirad, J.A. Wood, Y. Zhang, A.K. Ray, Multi-variable operational characteristic studies of on-column oxidative protein refolding at high loading concentrations, J. Chromatogr. A 1359 (2014) 70-75.

[19] P. Ricchiuto, A. V. Brukhno, S. Auer, Protein aggregation: Kinetics versus thermodynamics, J. Phys. Chem. B 116 (2012) 5384-5390.

[20] M. Mazzotti, G. Storti, M. Morbidelli, Optimal operation of simulated moving bed units for nonlinear chromatographic separations, J. Chromatogr. A (1997) 3–24.

[21] S. Pan, M. Zelger, R. Hahn, A. Jungbauer, Continuous protein refolding in a tubular reactor, Chem. Eng. Sci. 116 (2014) 763-772.

[22] F. Rodriguez- Ropero and N. van der Vegt, On the urea induced collapse of a water soluble polymer, Phys. Chem. Chem. Phys. (2015) DOI: 10.1039/C4CP05314A.

[23] Z. Xia, P. Das, E.I. Shakhnovich, R. Zhou, Collapse of unfolded proteins in a mixture of denaturants, J. Am. Chem. Soc. 134 (2012) 18266-18274.

[24] A.V.S. A. V. Volynskaya, S. A. Murasheva, A. Yu. Skripkin, A lysozyme unfolding mechanism, Dokl. Phys. Chem. 430 (2010) 29-32.

# Chapter 5

# 5 Summary and Future Work Recommendations

In spite of the advances made to date for expression of protein-therapeutics using *E. coli* host [1], the existing technologies to recover active and high-purity product still suffer from low refolding yields, volumetric productivities and high buffer consumption for commercial scale production which impose constraints in terms of time and cost [2]. This research work focused on various process development and optimization strategies to improve on above mentioned process performance indicators in size exclusion chromatography which has been extensively used at lab scale for protein refolding purpose. The outcome of this research is meant to be general, with broad application for cost effective high-throughput inclusion-body-based protein production.

Chapter 2 of this thesis presents a multivariable investigation of various parameters of the refolding buffer namely buffer pH, sodium chloride and L-arginine concentrations. It was illustrated that the benefit of using L-arginine is twofold as, in addition to preventing aggregation of lysozyme for loading concentrations of up to 40 mg/mL, it introduced the possibility of adjusting the buffer redox potential by controlling the pH to assist reformation of disulfide bridges without the need for costly redox couple chemicals. L-arginine also reduced the pore accessibility for small molecules of urea and DTT at higher protein concentrations indicating gradual removal of these agents which may be an additional mechanism by which this additive prevents aggregation in size exclusion chromatography. Further experimentation is however required to test these observations for other protein systems.

In Chapter 3, various methods were introduced for high-throughput measurements of mass transfer parameters and distribution coefficients of refolding kinetic species by single-column experiments in order to establish an experimentally-verified mathematical model. These methods included non-reactive native protein, model protein compounds and equilibrium experiments to extract characteristic information for native protein, early collapsed intermediates and native-like intermediates respectively. Although this was only demonstrated for the model protein (lysozyme), the same methods can be used for other oxidative and non-oxidative refolding cases. It was also demonstrated that the

above mentioned parameters from size exclusion chromatography may be improperly estimated due to non-specific interaction of species with the chromatography gel even when L-arginine was used; which is commonly suggested for reduced protein interaction with surfaces and increased recovery in various chromatography methods [3–5]. In addition, the importance of kinetic studies in deciding on an appropriate refolding additive was discussed. In this work, L-arginine additive was found to prevent aggregation without compromising the speed of lysozyme refolding.

Finally, investigations on the suitability of utilizing single-column data towards prediction of the behavior of the protein in multi-column continuous simulated moving bed size exclusion chromatography (SMB-SEC) and design/operation of a continuous process were carried out; presented in Chapter 4 of this thesis. The SMB-SEC technology offers many advantages compared to single-column processing, including increased productivity per unit mass of solid phase, lower solvent consumption, and less diluted products all of which makes it an attractive technology for future industrial applications. Since operation of a SMB-SEC unit includes more degrees of freedom compared to single-column operation, a systematic optimization is essential to obtain the operating parameter setting (s) that result in optimum process performance indicators. Accordingly, a robust mathematical model is a necessary tool for successful operation of a SMB-SEC. Therefore, the model that was developed in Chapter 3 was modified to extend its applicability for SMB-SEC. The prediction of unfolded protein and urea concentrations through a SMB-SEC showed lower dilution factors compared to single-column operation in this work. As a lower dilution factor results in higher local protein concentrations, protein aggregation was introduced into the model by introduction of a critical concentration (i.e. solubility of early intermediates) above which aggregates are formed with second-order rate kinetics of aggregation. Lower dilution factor also translates to a higher denaturant local concentration, influencing measured kinetic and thermodynamic parameters by single-column experiments.

Therefore, a model with constant kinetic and thermodynamic parameters may not result in accurate prediction of native to early intermediate protein ratio and an under-prediction of total solubilized protein concentration at the product line. However, if under various

operation settings no or very low concentration of denaturant exist in the product line, as it was the case in this work based on simulations, the refolding reaction will continue off-column to recover protein as functional native protein. Consequently, the process performance indicators were defined based on total solubilized protein at the product outlet (early intermediates and native protein). As a result the model with constant kinetic and thermodynamic parameters can be used to predict the solubilized protein recovery and find the operational space that meets an appropriate criterion for recovery where in-column aggregation and precipitation is mainly prevented and productivity and buffer consumption criteria are optimized. Experimentation on SMB-SEC systems is still required in order to evaluate the validity of this approach under various operating conditions.

## 5.1 References


[1] C.J. Huang, H. Lin, X. Yang, Industrial production of recombinant therapeutics in Escherichia coli and its recent advancements, J. Ind. Microbiol. Biotechnol. 39 (2012) 383-399.

[2] E.J. Freydell, L.A. van der Wielen, M.H. Eppink, M. Ottens, Techno-economic evaluation of an inclusion body solubilization and recombinant protein refolding process, Biotechnol. Prog. 27 (2011) 1315-1328.

[3] Y. Shikiya, S. Tomita, T. Arakawa, K. Shiraki, Arginine inhibits adsorption of proteins on polystyrene surface, PLoS One 8 (2013) e70762.

[4] T. Arakawa, K. Tsumoto, K. Nagase, D. Ejima, The effects of arginine on protein binding and elution in hydrophobic interaction and ion-exchange chromatography, Protein Expr. Purif. 54 (2007) 110-116.

[5] D. Ejima, R. Yumioka, T. Arakawa, K. Tsumoto, Arginine as an effective additive in gel permeation chromatography, J. Chromatogr. A 1094 (2005) 49-55.

# Appendices

# Appendix A: MATLAB Codes

## A.1 Parameter fitting for model without aggregation

```
clear all
clc
format short e

x0=0.1 ;    % initial guess-refolding reaction rate constant (min-1)
lb=0;  % lower bound
ub=1;  % upper bound

options = optimset('Display','iter');

[x,fval] = fmincon(@functioncal,x0,[],[],[],[],lb,ub,[],options);
% Returns the value of the objective function functioncal at the
solution x

function f=functioncal(x)

global NPLATE
global N
global CfI
global tpulse

%-------------------------------------------------------------------------
NPLATE = 100; % Number of plates
N  = 4 * NPLATE; % Number of unknowns

%-------------------------------------------------------------------------
                    % transferring experimental data

CexpN=[0.00 0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
```

```
0.01    0.00    0.00    0.01
0.03    0.02    0.00    0.02
0.06    0.07    0.05    0.02
0.09    0.12    0.09    0.00
0.08    0.17    0.08    0.00
0.10    0.22    0.13    0.06
0.13    0.27    0.17    0.19
0.19    0.33    0.20    0.23
0.20    0.45    0.26    0.35
0.25    0.51    0.35    0.44
0.27    0.65    0.37    0.53
0.31    0.66    0.41    0.62
0.35    0.73    0.43    0.68
0.41    0.82    0.47    0.74
0.42    0.88    0.49    0.76
0.45    0.90    0.51    0.78
0.34    0.77    0.48    0.70
0.19    0.49    0.37    0.57
0.15    0.36    0.23    0.44
0.08    0.26    0.15    0.33
0.08    0.22    0.12    0.23
0.08    0.17    0.09    0.24
0.06    0.19    0.09    0.22
0.07    0.15    0.07    0.17
0.05    0.14    0.06    0.14
0.06    0.12    0.03    0.15
0.07    0.10    0.02    0.14
0.05    0.09    0.02    0.10
0.03    0.08    0.01    0.10
0.05    0.07    0.00    0.09
0.05    0.05    0.00    0.08
0.03    0.02    0.00    0.00
0.01    0.00    0.00    0.00
0.02    0.00    0.00    0.00
0.02    0.00    0.00    0.00
0.01    0.00    0.00    0.00
0.01    0.00    0.00    0.00
0.01    0.00    0.00    0.00
0.01    0.00    0.00    0.00
0.01    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00
];

%-------------------------------------------------------------------------
                            %ODE integration

t0 = 0;
tf= 60;
y0 = zeros(N,1);

tspan= linspace (t0,tf,61);
f = 0;

%--------------------------------
```

```
CfIm = [5 10 10 20]; % loading concentration (mg cm-3)
tpulsem = [1 1 0.5 0.5]; % Injection time (min)

for j=1:4

    CfI = CfIm (j);
    tpulse = tpulsem (j);

    [T, Y]=ode15s(@(t,y)elution_profile(t,y,x),tspan,y0);

%-------------------------------------------------------------------------
                    % calculating objective function

Cexp= CexpN (:,j);

for i=1:length(T)

    f=f+(Y(i,3*N/4)-Cexp(i))^2 ;

end

end


function dydt = elution_profile(t, y, x)

global NPLATE
global N
global CfI
global tpulse

%   --------------- pre-allocating the variables

L = 27; % Column length (cm)

h = L / (NPLATE - 1); % Computing the step size

AREA = 2; % Column cross-section (cm2)

E = 0.34; % Bed void volume fraction (-)

Q = 1; % Flow rate (mL/min)

U = Q / (E * AREA); % Velocity (cm min-1)

P = (1 - E) / E; % Phase ratio

DL = 10 * U * 17e-4 * E; % Dispersion coefficient (cm2 min-1)

KsecN = 0.44; % Native protein distribution coefficient (-)

KsecI = 0.08; % Early intermediate protein distribution coefficient (-)
```

```
KovI = 2.69; % Early intermediate overal mass transfer coefficient
(min-1)

KovN = 17.29; % Native protein overal mass transfer coefficient (min-1)

Krxn = x(1,1); % Refolding reaction rate constant (min-1)

%-------------------------------------------------------------------------
                        %Constants used in function

B1 = U / (2* h);
C1 = DL / (h ^2);
B2= U/ (12 * h);
C2=DL / (12 * (h ^2));

dydt = zeros(N,1);
KZF = lt(t, tpulse);
CfN = 0;
%-------------------------------------------------------------------------
----
%                         Defining the ODEs


dydt(1) =  C1 * ( KZF * CfI - 2 * y(1) +  y(2)) - B1 * (-KZF * CfI +
y(2)) - P * KovI * (KsecI * y(1) - y(N/4+1)) -Krxn* y(1);


dydt(N/4+1) = KovI * ( KsecI * y(1) - y(N/4+1)) - Krxn * y(N/4+1);


dydt(N/2+1) = C1 * (KZF * CfN - 2 * y(N/2+1) +  y(N/2+2)) - B1 * (- KZF
* CfN + y(N/2+2))- P * KovN * (KsecN * y(N/2+1) - y(3*N/4+1)) + Krxn *
y(1);


dydt(3*N/4+1) = KovN *( KsecN * y(N/2+1) - y(3*N/4+1)) + Krxn *
y(N/4+1);


dydt(2) = C2 * ( - KZF * CfI + 16 * y(1) - 30 * y(2) + 16 * y(3) -
y(4)) - B2 * (KZF * CfI - 8 * y(1) + 8 * y(3) - y(4)) - P * KovI *
(KsecI * y(2) - y(N/4+2)) - Krxn * y(2);


dydt(N/4+2) = KovI * ( KsecI * y(2) - y(N/4+2)) - Krxn * y(N/4+2);


dydt(N/2+2) = C2 * (-KZF * CfN + 16 * y(N/2+1) - 30 * y(N/2+2) + 16 *
y(N/2+3) - y(N/2+4)) - B2 * (KZF * CfN - 8 * y(N/2+1) + 8 * y(N/2+3) -
y(N/2+4))- P * KovN * (KsecN * y(N/2+2) - y(3*N/4+2)) + Krxn * y(2);
```

```
dydt(3*N/4+2) = KovN *( KsecN * y(N/2+2) - y(3*N/4+2)) + Krxn *
y(N/4+2);


for I=3:N/4-2

dydt(I) = C2 * (- y (I-2) + 16 * y(I-1) - 30 * y(I) + 16 * y(I+1) -
y(I+2) )- B2 * ( y(I-2) - 8 * y(I-1) + 8 * y(I+1) - y(I+2)) - P * KovI
* (KsecI * y(I) - y(N/4+I)) - Krxn * y(I);


dydt(N/4+I) = KovI *( KsecI * y(I) - y(N/4+I)) - Krxn * y(N/4+I);


dydt(N/2+I) = C2 * (- y (N/2+I-2) + 16 * y(N/2+I-1) - 30 * y(N/2+I) +
16 * y(N/2+I+1) - y(N/2+I+2) ) - B2 * (y(N/2+I-2) - 8 * y(N/2+I-1) + 8
* y(N/2+I+1) - y(N/2+I+2)) - P * KovN * (KsecN * y(N/2+I) - y(3*N/4+I))
+ Krxn * y(I);


dydt(3*N/4+I) = KovN *( KsecN * y(N/2+I) -y(3*N/4+I)) + Krxn *
y(N/4+I);

end



dydt(N/4-1) = C2 * (- y (N/4-3) + 16 * y(N/4-2) - 30 * y(N/4-1) + 16 *
y(N/4) - y(N/4)) - B2 * (y(N/4-3) - 8 * y(N/4-2) + 8 * y(N/4) - y(N/4))
- P * KovI * (KsecI * y(N/4-1) - y(N/2-1)) - Krxn * y(N/4-1);


dydt(N/2-1) = KovI *( KsecI * y(N/4-1)-y(N/2-1)) - Krxn * y(N/2-1);


dydt(3*N/4-1) = C2 * (- y (3*N/4-3) + 16 * y(3*N/4-2) - 30 * y(3*N/4-1)
+ 16 * y(3*N/4) - y(3*N/4)) - B2 * (y(3*N/4-3) - 8 * y(3*N/4-2) + 8 *
y(3*N/4) - y(3*N/4)) - P * KovN * (KsecN * y(3*N/4-1) - y(N-1)) + Krxn
* y(N/4-1);


dydt(N-1) = KovN *( KsecN * y(3*N/4-1) -y(N-1)) + Krxn * y(N/2-1);


dydt(N/4) = C1 * (y(N/4-1) -2 * y(N/4) +  y(N/4)) - B1 * (- y(N/4-1) +
y(N/4)) - P * KovI * (KsecI * y(N/4) - y(N/2)) - Krxn * y(N/4);


dydt(N/2) = KovI *( KsecI * y(N/4)-y(N/2))- Krxn * y(N/2);
```

```
dydt(3*N/4) = C1 * (y(3*N/4-1) - 2 * y(3*N/4) + y(3*N/4)) - B1 * (-
y(3*N/4-1) + y(3*N/4)) - P * KovN *(KsecN * y(3*N/4) - y(N)) + Krxn *
y(N/4);

dydt(N) = KovN *( KsecN * y(3*N/4) -y(N)) + Krxn * y(N/2);
```

## A. 2 Parameter fitting for model with aggregation and Gaussian injection profile

```
clear all
clc
format short e

x0=0.1 ;  % initial guess-aggregation reaction rate constant (min-1)

options = optimset('Display','iter');

[x,fval] = fminsearch(@functioncal,x0,options);

function f=functioncal(x)

global NPLATE
global N
global a
%-------------------------------------------------------------------------
-----------------------------------

NPLATE = 100;      % Number of plates
N  = 4 * NPLATE;  % Number of unknowns

 %-------------------------------------------------------------------------
                   % transferring experimental data

CexpN=[0.00 0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.01    0.00    0.00    0.01    0.02
0.01    0.00    0.03    0.02    0.00    0.02    0.04
```

```
0.01    0.01    0.06    0.07    0.05    0.02    0.04
0.07    0.04    0.09    0.12    0.09    0.00    0.09
0.09    0.07    0.08    0.17    0.08    0.00    0.09
0.10    0.11    0.10    0.22    0.13    0.06    0.13
0.19    0.18    0.13    0.27    0.17    0.19    0.20
0.26    0.26    0.19    0.33    0.20    0.23    0.22
0.36    0.37    0.20    0.45    0.26    0.35    0.41
0.40    0.47    0.25    0.51    0.35    0.44    0.41
0.57    0.58    0.27    0.65    0.37    0.53    0.54
0.68    0.64    0.31    0.66    0.41    0.62    0.67
0.80    0.78    0.35    0.73    0.43    0.68    0.89
0.92    0.86    0.41    0.82    0.47    0.74    0.95
1.11    0.95    0.42    0.88    0.49    0.76    0.99
1.33    0.99    0.45    0.90    0.51    0.78    1.22
1.29    0.80    0.34    0.77    0.48    0.70    1.25
1.06    0.59    0.19    0.49    0.37    0.57    1.22
0.77    0.40    0.15    0.36    0.23    0.44    1.03
0.73    0.28    0.08    0.26    0.15    0.33    0.86
0.62    0.18    0.08    0.22    0.12    0.23    0.70
0.53    0.17    0.08    0.17    0.09    0.24    0.59
0.42    0.13    0.06    0.19    0.09    0.22    0.48
0.46    0.11    0.07    0.15    0.07    0.17    0.46
0.42    0.10    0.05    0.14    0.06    0.14    0.39
0.33    0.07    0.06    0.12    0.03    0.15    0.35
0.36    0.05    0.07    0.10    0.02    0.14    0.33
0.20    0.05    0.05    0.09    0.02    0.10    0.28
0.17    0.04    0.03    0.08    0.01    0.10    0.26
0.10    0.03    0.05    0.07    0.00    0.09    0.26
0.05    0.00    0.05    0.05    0.00    0.08    0.22
0.02    0.00    0.03    0.02    0.00    0.00    0.20
0.00    0.00    0.01    0.00    0.00    0.00    0.15
0.00    0.00    0.02    0.00    0.00    0.00    0.13
0.00    0.00    0.02    0.00    0.00    0.00    0.11
0.00    0.00    0.01    0.00    0.00    0.00    0.07
0.00    0.00    0.01    0.00    0.00    0.00    0.15
0.00    0.00    0.01    0.00    0.00    0.00    0.13
0.00    0.00    0.01    0.00    0.00    0.00    0.11
0.00    0.00    0.01    0.00    0.00    0.00    0.07
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00
];

 %-------------------------------------------------------------------
                            %ODE integration

t0 = 0;
tf= 60;
y0 = zeros(N,1);

tspan= linspace (t0,tf,61);
f = 0;

%--------------------------------
am = [26.56 13.28 3.36 6.59 3.36 6.59 16.6];
```

```
for j=1:7

    a = am (j);

    [T, Y]=ode15s(@(t,y)elution_profile(t,y,x),tspan,y0);

%-------------------------------------------------------------------------
                    % calculating objective function

Cexp= CexpN (:,j);

for i=1:length(T)

    f=f+(Y(i,3*N/4)-Cexp(i))^2 ;

end

end
function dydt = elution_profile(t, y, x)

global NPLATE
global N
global a


%   --------------- pre-allocating the variables

L = 27; % Column length (cm)

h = L / (NPLATE - 1); % Computing the step size

AREA = 2; % Column cross-section (cm2)

E = 0.34; % Bed void volume fraction (-)

Q = 1; % Flow rate (mL/min)

U = Q / (E * AREA); % Velocity (cm min-1)

P = (1 - E) / E; % Phase ratio

DL = 10 * U * 17e-4 * E; % Dispersion coefficient (cm2 min-1)

KsecN = 0.44; % Native protein distribution coefficient (-)

KsecI = 0.08; % Early intermediate protein distribution coefficient (-)

KovI = 2.69; % Early intermediate overal mass transfer coefficient
(min-1)

KovN = 17.29; % Native protein overal mass transfer coefficient (min-1)
```

```
Krxn = 0.08; % Refolding reaction rate constant (min-1)

kag = x(1,1); % Aggregation reaction rate constant (mLmin-1mg-1)

ys = 4.4; % Early intermediate solubility (mgmL-1)
%-------------------------------------------------------------------------
                    % Constants used in function

B1 = U / (2* h);
C1 = DL / (h ^2);
B2= U/ (12 * h);
C2=DL / (12 * (h ^2));

dydt = zeros(N,1);
CfN = 0;
             % Gaussian function and smooth step function

b = 1.5;
c = 0.58;
k = 100;
KZF = a * exp(-0.5*((t-b)/c)^2);
%-------------------------------------------------------------------------
%                       Defining the ODEs


dydt(1) =  C1 * ( KZF - 2 * y(1) +  y(2)) - B1 * (-KZF + y(2)) - P *
KovI * (KsecI * y(1) - y(N/4+1)) -Krxn* y(1)- 2*  kag * (1/(1+exp(-
2*k*(y(1)-ys))))*y(1)*y(1);


dydt(N/4+1) = KovI * ( KsecI * y(1) - y(N/4+1)) - Krxn * y(N/4+1)-2 *
(1/(1+exp(-2*k*(y(N/4+1)-ys))))* kag *(y(N/4+1))*(y(N/4+1));


dydt(N/2+1) = C1 * (KZF * CfN - 2 * y(N/2+1) +  y(N/2+2)) - B1 * (- KZF
* CfN + y(N/2+2))- P * KovN * (KsecN * y(N/2+1) - y(3*N/4+1)) + Krxn *
y(1);


dydt(3*N/4+1) = KovN *( KsecN * y(N/2+1) - y(3*N/4+1)) + Krxn *
y(N/4+1);


dydt(2) = C2 * ( - KZF + 16 * y(1) - 30 * y(2) + 16 * y(3) - y(4)) - B2
* (KZF - 8 * y(1) + 8 * y(3) - y(4)) - P * KovI * (KsecI * y(2) -
y(N/4+2)) - Krxn * y(2)-2* kag * (1/(1+exp(-2*k*(y(2)-
ys))))*(y(2))*(y(2));


dydt(N/4+2) = KovI * ( KsecI * y(2) - y(N/4+2)) - Krxn * y(N/4+2)-2 *
(1/(1+exp(-2*k*(y(N/4+2)-ys))))* kag *(y(N/4+2))*(y(N/4+2));
```

```
dydt(N/2+2) = C2 * (-KZF * CfN + 16 * y(N/2+1) - 30 * y(N/2+2) + 16 *
y(N/2+3) - y(N/2+4)) - B2 * (KZF * CfN - 8 * y(N/2+1) + 8 * y(N/2+3) -
y(N/2+4))- P * KovN * (KsecN * y(N/2+2) - y(3*N/4+2)) + Krxn * y(2);


dydt(3*N/4+2) = KovN *( KsecN * y(N/2+2) - y(3*N/4+2)) + Krxn *
y(N/4+2);


for I=3:N/4-2

dydt(I) = C2 * (- y (I-2) + 16 * y(I-1) - 30 * y(I) + 16 * y(I+1) -
y(I+2) )- B2 * ( y(I-2) - 8 * y(I-1) + 8 * y(I+1) - y(I+2)) - P * KovI
* (KsecI * y(I) - y(N/4+I)) - Krxn * y(I)-2 * (1/(1+exp(-2*k*(y(I)-
ys))))* kag *(y(I))*(y(I));


dydt(N/4+I) = KovI *( KsecI * y(I) - y(N/4+I)) - Krxn * y(N/4+I)-2 *
(1/(1+exp(-2*k*(y(N/4+I)-ys))))* kag *(y(N/4+I))*(y(N/4+I));


dydt(N/2+I) = C2 * (- y (N/2+I-2) + 16 * y(N/2+I-1) - 30 * y(N/2+I) +
16 * y(N/2+I+1) - y(N/2+I+2) ) - B2 * (y(N/2+I-2) - 8 * y(N/2+I-1) + 8
* y(N/2+I+1) - y(N/2+I+2)) - P * KovN * (KsecN * y(N/2+I) - y(3*N/4+I))
+ Krxn * y(I);


dydt(3*N/4+I) = KovN *( KsecN * y(N/2+I) -y(3*N/4+I)) + Krxn *
y(N/4+I);

end



dydt(N/4-1) = C2 * (- y (N/4-3) + 16 * y(N/4-2) - 30 * y(N/4-1) + 16 *
y(N/4) - y(N/4)) - B2 * (y(N/4-3) - 8 * y(N/4-2) + 8 * y(N/4) - y(N/4))
- P * KovI * (KsecI * y(N/4-1) - y(N/2-1)) - Krxn * y(N/4-1)-2*
(1/(1+exp(-2*k*(y(N/4-1)-ys))))* kag *(y(N/4-1))*(y(N/4-1));


dydt(N/2-1) = KovI *( KsecI * y(N/4-1)-y(N/2-1)) - Krxn * y(N/2-1)-2 *
(1/(1+exp(-2*k*(y(N/2-1)-ys))))* kag *(y(N/2-1))*(y(N/2-1));


dydt(3*N/4-1) = C2 * (- y (3*N/4-3) + 16 * y(3*N/4-2) - 30 * y(3*N/4-1)
+ 16 * y(3*N/4) - y(3*N/4)) - B2 * (y(3*N/4-3) - 8 * y(3*N/4-2) + 8 *
y(3*N/4) - y(3*N/4)) - P * KovN * (KsecN * y(3*N/4-1) - y(N-1)) + Krxn
* y(N/4-1);


dydt(N-1) = KovN *( KsecN * y(3*N/4-1) -y(N-1)) + Krxn * y(N/2-1);
```

```
dydt(N/4) = C1 * (y(N/4-1) -2 * y(N/4) +  y(N/4)) - B1 * (- y(N/4-1) +
y(N/4)) - P * KovI * (KsecI * y(N/4) - y(N/2)) - Krxn * y(N/4)- 2 *
(1/(1+exp(-2*k*(y(N/4)-ys))))* kag *(y(N/4))*(y(N/4));


dydt(N/2) = KovI *( KsecI * y(N/4)-y(N/2))- Krxn * y(N/2)-2 *
(1/(1+exp(-2*k*(y(N/2)-ys))))* kag *(y(N/2))*(y(N/2));


dydt(3*N/4) = C1 * (y(3*N/4-1) - 2 * y(3*N/4) + y(3*N/4)) - B1 * (-
y(3*N/4-1) + y(3*N/4)) - P * KovN *(KsecN * y(3*N/4) - y(N)) + Krxn *
y(N/4);

dydt(N) = KovN *( KsecN * y(3*N/4) -y(N)) + Krxn * y(N/2);
```

## A.3 SMB-SEC

```
clear all
clc
format long

global NP

global n1
global n2
global n3

global N

global QF

global QD
global Q2
global Q3
global Q4

global CfI
global CfN

NP = 50; % Number of plates for each column

n1 = 2; %Number of columns in section 1
n2 = 2; %Number of columns in section 2
n3 = 2; %Number of columns in section 3
n4 = 2; %Number of columns in section 4

Nt = n1+n2+n3+n4; %Total number of columns

Vc = 6; % Volume of each column (mL)

N = 4 * NP * Nt; % Number of unknowns
```

```
QD = 2.4; % Eluent flow rate (mL min-1)
QF = 0.26; % Feed flow rate (mL min-1)
QRa = 0.66; % Raffinate flow rate (mL min-1)
Q4 = 1.1; % Internal flow rate zone 4 and waste
Q3 = Q4 + QRa; % Internal flow rate zone 3
Q2 = Q4+QRa-QF; %Internal flow rate zone 2

CfI = 30; % Feed concentration mg mL-1
CfN = 0;

Nswitch = 30; % Number of switch times

%-------------------------------------------------------------------------
                            % ODE integration

y0 = zeros(N,1);
y1 = zeros(1,NP);
y2 = zeros(1,NP);
y3 = zeros(1,NP);
y4 = zeros(1,NP);
RaI =zeros (Nswitch,1);
RaN = zeros (Nswitch,1);
ExI = zeros (Nswitch,1);
ExN = zeros (Nswitch,1);
RecI = zeros (Nswitch,1);
RecN = zeros (Nswitch,1);
PrN = zeros (Nswitch,1);
BcN = zeros (Nswitch,1);

fid1 = fopen('profileI.txt','w');
fprintf(fid1,'%6s %6s %12s
%12s\r\n','k','j','Location','Intermediate');
fid2 = fopen('profileN.txt','w');
fprintf(fid2,'%6s %6s %12s %12s\r\n','k','j','Location','Native');

for k=1:Nswitch

t0 = 0;
ts= 3; % Switch time
tint = 101;

tspan= linspace(t0,ts,tint);

[T, Y] = ode15s(@SMB_Simulation, tspan, y0);


Yout = zeros (length(T),Nswitch*N);

Yout(:,(k-1)*N+1:k*N)= Y(:,1:N);


if (mod(k,8)== 0)

for j=1:10:tint %beginning, mid and end of each cycle
```

```
 for I=NP/10:NP/10:N/4; %every 5plate


fprintf(fid1,'%12d %12d %12d %4.2f\r\n',k,j,I,Yout(j, (k-1)*N +I)); % I
concentration profile along the SMB columns
fprintf(fid2,'%12d %12d %12d %4.2f\r\n',k,j,I,Yout(j, (k-1)*N +N/2+I));
% N concentration profile along the SMB columns

 end
 end

end


RaI(k,1) = (trapz(T,Yout(:,(k-1)* N + (n1+n2+n3)* NP)))/ ts; % Cyclic
steady state concentration of I in Raffinate stream
RaN(k,1) = (trapz(T,Yout(:,(k-1)* N + N/2+(n1+n2+n3)* NP)))/ ts;%
Cyclic steady state concentration of N in Raffinate stream
ExI(k,1) = (trapz(T,Yout(:,(k-1)* N+ n1* NP)))/ ts ;% Cyclic steady
state concentration of I in Extract stream
ExN(k,1) = (trapz(T,Yout(:,(k-1)* N+ N/2+n1* NP)))/ ts;% Cyclic steady
state concentration of N in Extract stream
WI(k,1) = (trapz(T,Yout(:,(k-1)* N + (n1+n2+n3+n4)* NP)))/ ts;% Cyclic
steady state concentration of I in Waste stream
WN(k,1) = (trapz(T,Yout(:,(k-1)* N + N/2+(n1+n2+n3+n4)* NP)))/ ts;%
Cyclic steady state concentration of N in Waste stream



RecN(k,1)=RaN(k,1)*(Q3-Q4)/(QF*CfI) ; %Recovery N
RecI(k,1)=RaI(k,1)*(Q3-Q4)/(QF*CfI);%Recovery I
PrN(k,1)= (RaN(k,1)+RaI(k,1))*(Q3-Q4)/(Nt*Vc);% Volumetric productivity
in terms of solubilized protein (mg mL-1 min-1)
BcN(k,1)= QD/ ((RaN(k,1)+RaI(k,1))*(Q3-Q4));% Buffer consumption (mL
mg-1)


for I=1:NP % store column 1 information @ tint

y1(1,I) = Y(tint,I);
y2(1,I) = Y(tint,N/4+I);
y3(1,I) = Y(tint,N/2+I);
y4(1,I) = Y(tint,3*N/4+I);

end

for J=0:Nt-2 % switching column in direction of solid movement

for I=1:NP
        Y(tint,I+J*NP) = Y(tint,I+(J+1)*NP);
        Y(tint,N/4+I+J*NP)= Y(tint,N/4+I+(J+1)*NP);
        Y(tint,N/2+I+J*NP)= Y(tint, N/2+I+(J+1)*NP);
        Y(tint,3*N/4+I+J*NP)= Y(tint,3*N/4+I+(J+1)*NP);
```

```
end
end

for I=1:NP
        Y(tint,N/4-NP+I)= y1(1,I);
        Y(tint,N/2-NP+I)= y2(1,I);
        Y(tint,3*N/4-NP+I)= y3(1,I);
        Y(tint,N-NP+I)= y4(1,I);
end

y0 = Y(tint,:).'; % Initial condition for new cycle


end

fclose(fid1);
fclose(fid2);


function dydt = SMB_Simulation(t,y)

global NP

global n1
global n2
global n3

global N

global QF

global QD
global Q2
global Q3
global Q4

global CfI
global CfN


%   --------------- pre-allocating the variables
AREA = 0.78; % Column cross section (cm2)
E = 0.34; % Bed void volume fraction (-)
L = 8; % Column length (cm)
KsecI = 0.08; % Equilibrium constant for Intermediate protein (-)
KsecN = 0.44; % Equilibrium constant for Native protein (-)
KovI = 2.69; % Mass transfer coefficient for Intermediate protein (min-
1)
KovN = 17.29; % Mass transfer coefficient for Native protein (min-1)
Krxn = 0.08; % Refolding reaction rate constant (min-1)
Kag = 0.08; % Aggregation reaction rate constant (min-1)

ys = 4.4; % Early intermediate solubility (mg mL-1)
k = 1000;
```

```
h =L /(NP - 1); % Axial step size
P = (1 - E) / E; % Phase ratio

%------------------------------------------------------------------------
PP = n1 * NP;
PQ = PP + n2 * NP;
PQR =PQ + n3 * NP;
%------------------------------------------------------------------------
dydt = zeros(N,1);
%------------------------------------------------------------------------

% Stage 1 &2 -------------Eluent point

U = QD/(AREA * E);
B1 = U/(2 * h);
Dl = 10 * U * 17e-4 * E;
C1 = Dl/(h ^ 2);
B2=U/(12 * h);
C2=Dl/(12 *h ^2);


% Central second order

dydt(1) = C1 * (0 - 2 * y(1) +  y(2)) - B1 * (-0 + y(2)) - P * KovI *
(KsecI * y(1) - y(N/4+1)) - Krxn * y(1)- 2*  Kag * (1/(1+exp(-
2*k*(y(1)-ys))))*y(1)*y(1);    % I liquid


dydt(N/4+1) = KovI * ( KsecI * y(1) - y(N/4+1)) - Krxn * y(N/4+1)-2 *
(1/(1+exp(-2*k*(y(N/4+1)-ys))))* Kag *(y(N/4+1))*(y(N/4+1));   % I
solid


dydt(N/2+1) = C1 * (0 - 2 * y(N/2+1) +  y(N/2+2)) - B1 * (- 0 +
y(N/2+2)) - P * KovN * (KsecN * y(N/2+1) - y(3*N/4+1)) + Krxn * y(1);


dydt(3*N/4+1) = KovN *( KsecN * y(N/2+1) - y(3*N/4+1)) + Krxn *
y(N/4+1);

% Central fourth order

dydt(2) = C2 * ( - 0 + 16 * y(1) - 30 * y(2) + 16 * y(3) - y(4)) - B2 *
(0 - 8 * y(1) + 8 * y(3) - y(4)) - P * KovI * (KsecI * y(2) - y(N/4+2))
- Krxn * y(2)-2* Kag * (1/(1+exp(-2*k*(y(2)-ys))))*(y(2))*(y(2));



dydt(N/4+2) = KovI * ( KsecI * y(2) - y(N/4+2)) - Krxn * y(N/4+2)-2 *
(1/(1+exp(-2*k*(y(N/4+2)-ys))))* Kag *(y(N/4+2))*(y(N/4+2));
```

```
dydt(N/2+2) = C2 * (-0 + 16 * y(N/2+1) - 30 * y(N/2+2) + 16 * y(N/2+3)
- y(N/2+4)) - B2 * (0- 8 * y(N/2+1) + 8 * y(N/2+3) - y(N/2+4)) - P *
KovN * (KsecN * y(N/2+2) - y(3*N/4+2)) + Krxn * y(2);



dydt(3*N/4+2) = KovN *( KsecN * y(N/2+2) - y(3*N/4+2)) + Krxn *
y(N/4+2);


% Stage 3-PP --------Section 1

for I=3:PP

dydt(I) = C2 * (- y(I-2) + 16 * y(I-1) - 30 * y(I) + 16 * y(I+1) -
y(I+2) )- B2 * ( y(I-2) - 8 * y(I-1) + 8 * y(I+1) - y(I+2)) - P * KovI
* (KsecI * y(I) - y(N/4+I)) - Krxn * y(I)-2 * (1/(1+exp(-2*k*(y(I)-
ys))))* Kag *(y(I))*(y(I));


dydt(N/4+I) = KovI *( KsecI * y(I) - y(N/4+I)) - Krxn * y(N/4+I)-2 *
(1/(1+exp(-2*k*(y(N/4+I)-ys))))* Kag *(y(N/4+I))*(y(N/4+I));


dydt(N/2+I) = C2 * (- y(N/2+I-2) + 16 * y(N/2+I-1) - 30 * y(N/2+I) + 16
* y(N/2+I+1) - y(N/2+I+2) ) - B2 * (y(N/2+I-2) - 8 * y(N/2+I-1) + 8 *
y(N/2+I+1) - y(N/2+I+2)) - P * KovN * (KsecN * y(N/2+I) - y(3*N/4+I)) +
Krxn * y(I);


dydt(3*N/4+I) = KovN *( KsecN * y(N/2+I) - y(3*N/4+I)) + Krxn *
y(N/4+I);

 end

% Stage (PP+1)-(PP+2) ---------Section2

U = Q2 / (AREA * E);
B1 = U / (2 * h);
Dl = 10 * U * 17e-4 * E;
C1 = Dl / (h ^ 2);
B2= U / (12 * h);
C2=Dl / (12 * h ^2);


% Stage (PP+1)-(PQ-2) ---------Section2


 for I=PP+1:PQ-2

dydt(I) = C2 * (- y (I-2) + 16 * y(I-1) - 30 * y(I) + 16 * y(I+1) -
y(I+2) )- B2 * ( y(I-2) - 8 * y(I-1) + 8 * y(I+1) - y(I+2)) - P * KovI
```

```
* (KsecI * y(I) - y(N/4+I)) - Krxn * y(I)-2 * (1/(1+exp(-2*k*(y(I)-
ys))))* Kag *(y(I))*(y(I));


dydt(N/4+I) = KovI *( KsecI * y(I) - y(N/4+I)) - Krxn * y(N/4+I)-2 *
(1/(1+exp(-2*k*(y(N/4+I)-ys))))* Kag *(y(N/4+I))*(y(N/4+I));


dydt(N/2+I) = C2 * (- y(N/2+I-2) + 16 * y(N/2+I-1) - 30 * y(N/2+I) + 16
* y(N/2+I+1) - y(N/2+I+2) ) - B2 * (y(N/2+I-2) - 8 * y(N/2+I-1) + 8 *
y(N/2+I+1) - y(N/2+I+2)) - P * KovN * (KsecN * y(N/2+I) - y(3*N/4+I)) +
Krxn * y(I);


dydt(3*N/4+I) = KovN *( KsecN * y(N/2+I) - y(3*N/4+I)) + Krxn *
y(N/4+I);

 end

% Stage (PQ-1)-(PQ) ---------Section2

dydt(PQ-1) = C2 * (- y(PQ-3) + 16 * y(PQ-2) - 30 * y(PQ-1) + 16 * y(PQ)
- y(PQ)) - B2 * (y(PQ-3) - 8 * y(PQ-2) +  8 * y(PQ) - y(PQ)) - P * KovI
* (KsecI * y(PQ-1) - y(N/4 +PQ-1)) - Krxn * y(PQ-1)-2 * (1/(1+exp(-
2*k*(y(PQ-1)-ys))))* Kag *(y(PQ-1))*(y(PQ-1));


dydt(N/4+PQ-1) = KovI *( KsecI * y(PQ-1)-y(N/4+PQ-1)) - Krxn *
y(N/4+PQ-1)-2 * (1/(1+exp(-2*k*(y(N/4+PQ-1)-ys))))* Kag *(y(N/4+PQ-
1))*(y(N/4+PQ-1));


dydt(N/2+PQ-1) = C2 * (- y(N/2+PQ-3) + 16 * y(N/2+PQ-2) - 30 *
y(N/2+PQ-1) + 16 * y(N/2+PQ) - y(N/2+PQ)) - B2 * (y(N/2+PQ-3) - 8 *
y(N/2+PQ-2) + 8 * y(N/2+PQ) - y(N/2+PQ)) - P * KovN * (KsecN *
y(N/2+PQ-1) - y(3*N/4+PQ-1)) + Krxn * y(PQ-1);


dydt(3*N/4+PQ-1) = KovN *( KsecN * y(N/2+PQ-1) -y(3*N/4+PQ-1)) + Krxn *
y(N/4+PQ-1);


dydt(PQ) = C1 * (y(PQ-1) -2 * y(PQ) +  y(PQ)) - B1 * (- y(PQ-1) +
y(PQ)) - P * KovI * (KsecI * y(PQ) - y(N/4+PQ)) - Krxn * y(PQ)-2 *
(1/(1+exp(-2*k*(y(PQ)-ys))))* Kag *(y(PQ))*(y(PQ));


dydt(N/4+PQ) = KovI *( KsecI * y(PQ)-y(N/4+PQ)) - Krxn * y(N/4+PQ)-2 *
(1/(1+exp(-2*k*(y(N/4+PQ)-ys))))* Kag *(y(N/4+PQ))*(y(N/4+PQ));
```

```
dydt(N/2+PQ) = C1 * (y(N/2+PQ-1) - 2 * y(N/2+PQ) + y(N/2+PQ)) - B1 * (-
y(N/2+PQ-1) + y(N/2+PQ)) - P * KovN *(KsecN * y(N/2+PQ) - y(3*N/4+PQ))
+ Krxn * y(PQ);


dydt(3*N/4+PQ) = KovN *( KsecN * y(N/2+PQ) -y(3*N/4+PQ)) + Krxn *
y(N/4+PQ);



% Stage (PQ+1) AND (PQ+2)----------Feed point

U = Q3 / (AREA * E);
Dl = 10 * U * 17e-4 * E;
B1 = U / (2 * h);
C1 = Dl / (h ^ 2);
B2 = U / (12 * h);
C2 = Dl / (12 * h ^ 2);


dydt(PQ+1) = C1 * ((Q2 * y(PQ) / Q3 + QF * CfI/Q3) - 2 * y(PQ+1) +
y(PQ+2)) - B1 * (-(Q2 * y(PQ) / Q3 + QF * CfI / Q3) + y(PQ+2)) - P *
KovI * (KsecI * y(PQ+1) - y(N/4+PQ+1)) - Krxn * y(PQ+1)-2 * (1/(1+exp(-
2*k*(y(PQ+1)-ys))))* Kag *(y(PQ+1))*(y(PQ+1));


dydt(N/4+PQ+1) = KovI * ( KsecI * y(PQ+1) - y(N/4+PQ+1)) - Krxn *
y(N/4+PQ+1)-2 * (1/(1+exp(-2*k*(y(N/4+PQ+1)-ys))))* Kag
*(y(N/4+PQ+1))*(y(N/4+PQ+1));

dydt(N/2+PQ+1) = C1 * ((Q2 * y(N/2+PQ) / Q3 + QF * CfN / Q3) - 2 *
y(N/2+PQ+1) +  y(N/2+PQ+2)) - B1 * (-(Q2 * y(N/2+PQ) / Q3 + QF * CfN /
Q3) + y(N/2+PQ+2)) - P * KovN * (KsecN * y(N/2+PQ+1) - y(3*N/4+PQ+1)) +
Krxn * y(PQ+1);


dydt(3*N/4+PQ+1) = KovN *( KsecN * y(N/2+PQ+1) - y(3*N/4+PQ+1)) + Krxn
* y(N/4+PQ+1);


dydt(PQ+2) = C2 * ( -(Q2 * y(PQ) / Q3 + QF * CfI / Q3)+ 16 * y(PQ+1) -
30 * y(PQ+2) + 16 * y(PQ+3) - y(PQ+4)) - B2 * ((Q2*y(PQ)/Q3+QF*CfI/Q3)
- 8 * y(PQ+1) + 8 * y(PQ+3) - y(PQ+4)) - P * KovI * (KsecI * y(PQ+2) -
y(N/4+PQ+2)) - Krxn * y(PQ+2)-2 * (1/(1+exp(-2*k*(y(PQ+2)-ys))))* Kag
*(y(PQ+2))*(y(PQ+2));



dydt(N/4+PQ+2) = KovI * ( KsecI * y(PQ+2) - y(N/4+PQ+2)) - Krxn *
y(N/4+PQ+2)-2 * (1/(1+exp(-2*k*(y(N/4+PQ+2)-ys))))* Kag
*(y(N/4+PQ+2))*(y(N/4+PQ+2));
```

```
dydt(N/2+PQ+2) = C2 * (-(Q2 * y(N/2+PQ) / Q3 + QF * CfN / Q3) + 16 *
y(N/2+PQ+1) - 30 * y(N/2+PQ+2) + 16 * y(N/2+PQ+3) - y(N/2+PQ+4)) - B2 *
((Q2 * y(N/2+PQ) / Q3 + QF * CfN / Q3) - 8 * y(N/2+PQ+1) + 8 *
y(N/2+PQ+3) - y(N/2+PQ+4))- P * KovN * (KsecN * y(N/2+PQ+2) -
y(3*N/4+PQ+2)) + Krxn * y(PQ+2);


dydt(3*N/4+PQ+2) = KovN *( KsecN * y(N/2+PQ+2) - y(3*N/4+PQ+2)) + Krxn
* y(N/4+PQ+2);


% Stage (PQ+3)-(PQR) -------Section 3

for I=PQ+3:PQR

dydt(I) = C2 * (- y (I-2) + 16 * y(I-1) - 30 * y(I) + 16 * y(I+1) -
y(I+2) )- B2 * ( y(I-2) - 8 * y(I-1) + 8 * y(I+1) - y(I+2)) - P * KovI
* (KsecI * y(I) - y(N/4+I)) - Krxn * y(I)-2 * (1/(1+exp(-2*k*(y(I)-
ys))))* Kag *(y(I))*(y(I));


dydt(N/4+I) = KovI *( KsecI * y(I) - y(N/4+I)) - Krxn * y(N/4+I)-2 *
(1/(1+exp(-2*k*(y(N/4+I)-ys))))* Kag *(y(N/4+I))*(y(N/4+I));


dydt(N/2+I) = C2 * (- y(N/2+I-2) + 16 * y(N/2+I-1) - 30 * y(N/2+I) + 16
* y(N/2+I+1) - y(N/2+I+2) ) - B2 * (y(N/2+I-2) - 8 * y(N/2+I-1) + 8 *
y(N/2+I+1) - y(N/2+I+2)) - P * KovN * (KsecN * y(N/2+I) - y(3*N/4+I)) +
Krxn * y(I);


dydt(3*N/4+I) = KovN *( KsecN * y(N/2+I) - y(3*N/4+I)) + Krxn *
y(N/4+I);

end

% Stage (PQR+1)-N/4-2-------section 4

U = Q4 / (AREA * E);
Dl = 10 * U * 17e-4 * E;
B1 = U / (2 * h);
C1 = Dl / (h ^ 2);
B2 = U / (12 * h);
C2 = Dl / (12 * h ^ 2);

for I = PQR+1:N/4-2

 dydt(I) = C2 * (- y (I-2) + 16 * y(I-1) - 30 * y(I) + 16 * y(I+1) -
y(I+2)) - B2 * ( y(I-2) - 8 * y(I-1) + 8 * y(I+1) - y(I+2)) - P * KovI
* (KsecI * y(I) - y(N/4+I)) - Krxn * y(I)-2 * (1/(1+exp(-2*k*(y(I)-
ys))))* Kag *(y(I))*(y(I));
```

```
dydt(N/4+I) = KovI *( KsecI * y(I) - y(N/4+I)) - Krxn * y(N/4+I)-2 *
(1/(1+exp(-2*k*(y(N/4+I)-ys))))* Kag *(y(N/4+I))*(y(N/4+I));


dydt(N/2+I) = C2 * (- y(N/2+I-2) + 16 * y(N/2+I-1) - 30 * y(N/2+I) + 16
* y(N/2+I+1) - y(N/2+I+2) ) - B2 * (y(N/2+I-2) - 8 * y(N/2+I-1) + 8 *
y(N/2+I+1) - y(N/2+I+2)) - P * KovN * (KsecN * y(N/2+I) - y(3*N/4+I)) +
Krxn * y(I);


dydt(3*N/4+I) = KovN *( KsecN * y(N/2+I) - y(3*N/4+I)) + Krxn *
y(N/4+I);

 end

% last stage ----------Section 4

dydt(N/4-1) = C2 * (- y(N/4-3) + 16 * y(N/4-2) - 30 * y(N/4-1) + 16 *
y(N/4) - y(N/4)) - B2 * (y(N/4-3) - 8 * y(N/4-2) + 8 * y(N/4) - y(N/4))
- P * KovI * (KsecI * y(N/4-1) - y(N/2-1)) - Krxn * y(N/4-1)-2*
(1/(1+exp(-2*k*(y(N/4-1)-ys))))* Kag *(y(N/4-1))*(y(N/4-1));


dydt(N/2-1) = KovI *( KsecI * y(N/4-1) - y(N/2-1)) - Krxn * y(N/2-1)-2
* (1/(1+exp(-2*k*(y(N/2-1)-ys))))* Kag *(y(N/2-1))*(y(N/2-1));


dydt(3*N/4-1) = C2 * (- y(3*N/4-3) + 16 * y(3*N/4-2) - 30 * y(3*N/4-1)
+ 16 * y(3*N/4) - y(3*N/4)) - B2 * (y(3*N/4-3) - 8 * y(3*N/4-2) + 8 *
y(3*N/4) - y(3*N/4)) - P * KovN * (KsecN * y(3*N/4-1) - y(N-1)) + Krxn
* y(N/4-1);


dydt(N-1) = KovN *( KsecN * y(3*N/4-1) -y(N-1)) + Krxn * y(N/2-1);


dydt(N/4) = C1 * (y(N/4-1) -2 * y(N/4) +  y(N/4)) - B1 * (- y(N/4-1) +
y(N/4)) - P * KovI * (KsecI * y(N/4) - y(N/2)) - Krxn * y(N/4)- 2 *
(1/(1+exp(-2*k*(y(N/4)-ys))))* Kag *(y(N/4))*(y(N/4));



dydt(N/2) = KovI *( KsecI * y(N/4) - y(N/2)) - Krxn * y(N/2)-2 *
(1/(1+exp(-2*k*(y(N/2)-ys))))* Kag *(y(N/2))*(y(N/2));

dydt(3*N/4) = C1 * (y(3*N/4-1) - 2 * y(3*N/4) + y(3*N/4)) - B1 * (-
y(3*N/4-1) + y(3*N/4)) - P * KovN *(KsecN * y(3*N/4) - y(N)) + Krxn *
y(N/4);


dydt(N) = KovN *( KsecN * y(3*N/4) - y(N)) + Krxn * y(N/2);
```

# Curriculum Vitae

| | |
|---|---|
| **Name:** | PEGAH SAREMIRAD |
| **Post-secondary Education and Degrees:** | Western University<br>London, Ontario, Canada<br>2010-2015 Ph.D.<br><br>Western University<br>London, Ontario, Canada<br>2008-2010 M.E.Sc<br><br>Sharif University of Technology<br>Tehran, Iran<br>2001-2006 B.Sc |
| **Honours and Awards:** | Province of Ontario Graduate Scholarship<br>2014-2015<br><br>Graduate Student Teaching Award (SOGS)<br>2014<br><br>Excellence in Teaching Award (Chemical Eng. Dept.)<br>2014 |
| **Related Work Experience** | Research Associate<br>Western University<br>2008-2015<br><br>Teaching Assistant<br>Western University<br>2008-2015 |

**Publications and Conferences:**

P. Saremirad, J.A. Wood, Y. Zhang, A.K. Ray, submitted to J. Chromatogr. A. 2015.
P. Saremirad, J.A. Wood, Y. Zhang, A.K. Ray, J. Chromatogr. A 2014, 1370, 145-155.
P. Saremirad, J.A. Wood, Y. Zhang, A.K. Ray, J. Chromatogr. A 2014, 1359, 70–75.
P. Saremirad, Y. Zhang, A.K. Ray, AIChE Annual Meeting, Nov 3-8, 2013, San Francisco, CA , USA, Oral Presentation.
P. Saremirad, H.G. Gomaa, J. Zhu, J. Mem. Sci. 2012, 405-406, 158-166.